UNIVERSITE PARIS-X-NANTERRE

UFR LLPH

DEPARTEMENT PHILOSOPHIE

ANNEE 1990-1991

MEMOIRE DE MAITRISE

PAR M. BERNARD PITOU

APPROCHES SUBJECTIVES ET OBJECTIVES DU CONCEPT DE PROBABILITE.

SOUS LA DIRECTION DE M. BERNARD RUYER ET DE MME. ANNE FAGOT-LARGEAULT

Table des matières

INTRODUCTION..... 3

CHAPITRE I, PROBABILISME ET PROBABILITE : EMERGENCE DE LA PROBABILITE COMME THEME PHILOSOPHIQUE. LIEN AVEC LE SCEPTICISME..... 15

CHAPITRE II, CHANCES ET PROBABILITES..... 22

CHAPITRE III, PROBABILITES SUBJECTIVES ET OBJECTIVES CHEZ LAPLACE..... 29

CHAPITRE IV FREQUENCES ET PROBABILITES..... 47

CHAPITRE V PROBABILITES PERSONNELLES..... 59

CONCLUSION..... 68

REFERENCES BIBLIOGRAPHIQUES..... 71

INTRODUCTION

D'où vient l'intérêt porté de nos jours par de nombreux philosophes de langue anglaise à la notion de probabilité ? Il me semble que cela tient à plusieurs raisons : la première est due au développement brusque des statistiques au XX^{e} siècle.

De nos jours les statistiques ont envahi toutes les sciences sociales, la psychologie, la sociologie, la médecine, les assurances, la biologie, l'agriculture, l'économétrie...etc. On peut parler de véritable révolution, qui transforme les critères et les moyens d'admission de la preuve scientifique.

En effet, jusqu'au milieu du XIX^{e} siècle, la méthode scientifique était fondée sur un schéma déductif inspiré des mathématiques où les preuves étaient déduites de manière rigoureuse de prémisses admises comme évidentes.

Or avec le développement de l'empirisme, le procès de la preuve se renverse et l'objectivité se retrouve du côté de l'expérience. Il faut partir désormais de l'expérience pour fonder, c'est-à-dire confirmer ou infirmer l'objectivité d'une hypothèse.

Ce renversement a particulièrement intéressé les philosophes, qui ont vu dans la statistique naissante une réponse au problème devenu classique depuis Hume de l'induction. Ceci d'autant plus qu'ils ont vu des statisticiens comme Fisher apporter des éléments au débat. D'autre part, toujours liés à ce problème de l'induction, on voit se développer des systèmes de logique inductive qui doivent, sur le plan formel cette fois-ci, créer une alternative à la logique déductive.

La troisième raison apparaît enfin liée à l'importance des probabilités dans la physique moderne, car, dans ce domaine la description probabiliste a semblé fournir une alternative au déterminisme.

Des savants, comme Einstein (dont la sentence « Dieu ne joue pas aux dés », est passée à la postérité) et des épistémologues comme Popper sont partis du principe qu'une description probabiliste des phénomènes physiques impliquait la présence d'un hasard objectif dans les choses, niant par là le déterminisme, à condition toutefois que ces probabilités fussent interprétées objectivement, et non pas comme l'effet d'une connaissance incomplète.

Cette double direction de l'extension de « l'empire de la chance » semble avoir comme corollaire, sur le plan épistémologique, une nouvelle émergence du probabilisme.

§2. Nous devons discuter en quelques mots de ces problèmes philosophiques. Le problème de l'induction peut se formuler ainsi. Comment pouvons-nous passer de l'observation de cas singuliers à l'affirmation de lois universelles, de telle manière que nos inférences soient valides ? Ou de manière plus spécifique : comment nos données empiriques peuvent conforter notre hypothèse théorique ?

La réponse de Hume à la première question est négative. Nous ne pouvons pas par induction conclure à une loi générale qui soit certaine. Les systèmes de logique inductive, ainsi que la statistique conçue selon Fisher pensent apporter une réponse à la seconde question. Si nous ne pouvons atteindre à la certitude dans nos inférences inductives, peut-être pourrions-nous donner un certain degré de probabilité à nos conclusions : La controverse sur les interprétations, subjectives ou objectives, du concept de probabilité nait dans cet arrière-plan.

Les uns, fréquentistes ou objectivistes veulent bannir la subjectivité de la science (Cf. « Pour une épistémologie sans sujet connaissant », de K. Popper) ; les

autres, appelés bayésiens, ou encore personnalistes, veulent au contraire que soient prises en compte dans les jugements les opinions, les croyances, et ce jusque dans des choix prétendument objectifs que les objectivistes pensent faire.

Par conséquent, il sera intéressant de les confronter plus tard.

§3. Le problème du déterminisme est tout aussi important. On peut définir le déterminisme comme la possibilité idéale pour la science de prévoir l'évolution d'un système physique de manière certaine à partir de données initiales spécifiées.

La science classique, en entendant par là la mécanique rationnelle, l'optique, et l'électromagnétisme, est déterministe, car à partir de la donnée des conditions initiales, qui sont les positions et les vitesses des particules matérielles pour la mécanique rationnelle, elle déduit les positions et les vitesses futures des particules.

Le concept central est alors celui de trajectoire, ensemble des points que décrit les particules. Si l'on remplace l'espace physique réel par l'espace des phases, la trajectoire devient une orbite, ou une courbe intégrale de l'espace des phases. Comment l'inférence est-elle possible ? On procède de deux manières : ou bien onapplique la deuxième loi de Newton qui dit que la résultante des forces appliquées au système est proportionne11e à l'accélération. On a alors un système d'équations différentielles du second ordre. Ou bien on considère certaines grandeurs invariantes telles que l'énergie ou la quantité de mouvement. On a alors des équations aux dérivées partielles qui lient coordonnées de positions et impulsions via l'Hami1tonien du système. Une fois données les conditions initiales, on peut déduire toutes les trajectoires possibles du système. Le déterminisme consiste alors dans cette possibilité idéale de spécifier les conditions initiales avec toute la précision voulue. Laplace exprime merveilleusement cette vue dans les premières pages de son *Essai philosophique sur les probabi1ités* :

« Nous devons donc envisager l'état présent de l'univers comme l'effet de son état antérieur, et comme la cause de celui qui va suivre. Une intelligence

qui pour un instant donné connaîtrait toutes les forces dont la nature est animée et la situation respective des êtres qui la composent, si d'ailleurs elle était assez vaste pour soumettre ces données à l'analyse, embrasserait dans la même formule les mouvements des plus grands corps de l'univers et ceux du plus léger atome : rien ne serait incertain pour elle, et l'avenir, comme le passé serait présent à ses yeux »[1]

Ce texte de Laplace dit plus sur le déterminisme que la définition que nous avons proposée. Il lie en effet déterminisme et causalité, le premier n'étant que l'expression de la seconde dans le cadre des théories physiques.

Cette identité est l'héritage du XVIII^{e} siècle dont on retrouve la trace jusqu'au XX^{e} siècle. C'est pourquoi, pour beaucoup de savants rationalistes, la négation du déterminisme est équivalente à la négation de la causalité. Parler de l'indéterminisme de la science ce serait introduire l'irrationalisme en sciences.

Bien que cette science classique (mais la mécanique statistique est-elle encore la science classique ?) ait connu des lois reposant sur des hypothèses probabilistes (loi de distribution des vitesses de Maxwell, distribution canonique de Boltzmann par exemple), la nécessité de telles lois était imputée aux faiblesses de l'esprit humain, à l'impossibilité pratique de résoudre les calculs.

§4. La remise en cause la plus radicale de ce déterminisme semble avoir été le fait de la Mécanique quantique.

Historiquement le problème qui a mené à la théorie des quanta est le problème du rayonnement du corps noir. Planck est conduit à introduire "sa" constante h pour montrer que l'énergie du rayonnement est quantifiée. L'idée est radicalement neuve et ressemble plus

[1] Laplace, *Essai philosophique sur les probabilités*, Christian Bourgeois éditeur, avec une préface de René Thom et une postface de Bernard Bru, Paris, 1986, p.32-33.

à une hypothèse ad hoc destinée à réconcilier la loi de Rayleigh-Jeans et la loi de Wien, et aurait fini dans la poussière des annales de physique, si quelques années après (1905) un jeune physicien (Albert Einstein) ne l'avait pas utilisée pour expliquer l'effet photo-électrique. L'hypothèse des quanta était ainsi introduite dans la science et allait connaître un succès fabuleux.

Cependant, elle n'allait pas être acceptée de suite par la communauté scientifique, car depuis le début du XIX^{e} siècle, on avait cru montrer le caractère ondulatoire de la lumière (expériences des fentes de Young). La lumière solaire diffractée à travers deux fentes produisait des phénomènes d'interférence. Or là où la physique classique voyait des interférences, elle les associait à la présence d'ondes.

Avec la théorie de l'effet photo-électrique, la dualité onde-corpuscule était soulignée pour la première fois.

D'autres faits venaient militer en faveur de la nouvelle théorie des quanta.

En 1911, au laboratoire Cavendish, Rutherford met en évidence la présence des noyaux atomiques par la diffusion des particules alpha. Il propose alors un modèle de l'atome (modèle planétaire) avec des électrons chargés négativement qui tournent autour d'un noyau chargé positivement. Ces électrons sont régis par la loi d'interaction coulombienne, c'est-à-dire par une loi de la physique classique. Mais son modèle est inadéquat pour expliquer la stabilité des atomes, car, selon ces mêmes lois de la physique classique, les électrons devraient tomber sur le noyau au bout d'un temps très court.

Bohr, qui travaille à son laboratoire, émet alors en 1913 l'hypothèse que les électrons se déplacent sur des orbites particulières et bien définies par leur niveau d'énergie premier postulat). Et ils ne peuvent passer d'un niveau à l'autre qu'en absorbant ou en émettant de l'énergie de manière discrète (deuxième et troisième postulats). Ainsi se trouve expliqué le phénomène des raies spectrales émises par les corps chimiques. On remarque que dans cette ancienne théorie des quanta le

concept d'orbite ou de trajectoire de l'électron autour du noyau garde un contenu intuitif et réel.

La théorie de Bohr avait cependant quelques défauts, car elle ne fournissait pas d'explication sur l'intensité des raies observées.

La dernière étape est 1924 : Louis de Broglie introduit la notion d'ondes matérielles associées à des particules de matière par la relation $\lambda=h/mv$ (λ longueur d'onde ; mv quantité de mouvement, masse multipliée par la vitesse). La relation de de Broglie est vérifiée expérimentalement lorsqu'on 1927 Davisson et Germer réalisent la diffraction des électrons par un cristal de nickel, qui est l'analogue de la diffraction de la lumière par un réseau cristallin.

En 1927 le temps est mûr pour une synthèse théorique. Elle se fait autour des grands noms : Bohr, Heisenberg. Pauli, Born, Schrödinger, de Broglie, Dirac.

Résumons-nous, et montrons l'introduction des concepts probabilistes dans la nouvelle mécanique.

On avait cru la lumière de nature ondulatoire. L'effet photo-électrique ne peut s'expliquer qu'en introduisant l'hypothèse de "quanta de lumière". De l'autre côté on avait admis jusqu'en 1924 que les électrons étaient des particules. Or voilà qu'on réalise des expériences de diffraction des électrons. En 1927, le fait était acquis : les objets quantiques se comportaient à la fois comme des ondes et comme des particules.

Que devient la notion de trajectoire que la physique classique associait à la particule ? Cette trajectoire est complètement déterminée par la donnée des positions et des vitesses initiales de la particule, et par le système d'équations différentielles qui relie ces positions et ces vitesses. On suppose de plus que les positions et les vitesses sont des grandeurs indépendantes. Mais Heisenberg montre que ces grandeurs ne sont pas indépendantes en MQ, car les opérateurs qui les représentent ne sont pas commutables[2]. Il déduit alors les relations d'incertitude. On

[2] « Les principes physiques de la théorie des quanta », trad. française de « Die Physikalischen Prinzipien der Quantentheorie » (1930), Paris, 1972, nouveau tirage, p.15.

ne peut déterminer en même temps la position et la vitesse d'une particule parce qu'elles sont des grandeurs conjuguées (Cf. Algèbre non commutative d'Alain Connes)

On illustre généralement cette indétermination par l'expérience suivante : on projette un faisceau monocinétique d'électrons à travers deux fentes minces, dont la largeur est de l'ordre de grandeur de la longueur d'ondes des électrons. On obtient alors des figures d'interférences. Si l'on adopte un point de vue corpusculaire, et si l'on veut déterminer par quelle fente est passé l'électron en plaçant juste derrière elles une source lumineuse qui "éclairerait" les électrons, les figures d'interférence sont alors détruites[3]. L'expérience est interprétée comme suit : pour voir un électron, c'est-à-dire pour déterminer sa position il doit heurter un photon. Or ce dernier a une quantité de mouvement déterminée par les relations d'Einstein-de Broglie. Mais par l'effet Compton, le photon va communiquer cette quantité de mouvement à l'électron. Par conséquent, après le choc, la quantité de mouvement de l'électron ne sera plus la même. Il y aura incertitude sur cette grandeur[4].

L'exigence de précision parfaite sur les conditions initiales devient une illusion. Par conséquent le déterminisme, qui était fondé sur cette possibilité, est écarté de la science. Le corollaire de cet indéterminisme est l'introduction d'une description statistique fondée sur des probabilités.

Philosophiquement, le projet de Heisenberg est inspiré par l'idée positiviste suivante, qui vient de Bohr : éliminer de la nouvelle théorie des quanta les grandeurs qui ne sont pas observables. Or, dans la nouvelle mécanique les seules grandeurs observables sont les fréquences et les niveaux d'énergie. A chaque grandeur observable, la nouvelle mécanique associe un opérateur linéaire hermitien (de Charles Hermite, mathématicien français).

---

[3] Cf. Feynman, Leighton, Sands, *Lectures on Physics*, T.III, chap.1

[4] Cf. pour plus de détails, la discussion des expériences in Heisenberg, op.cit. p.15 sqq.

Cet opérateur appliqué à une fonction qui représente l'état du système (fonction d'onde ou vecteur d'état) ne peut prendre que certaines valeurs appelées valeurs propres. Les résultats de la mesure (ou de l'observation) sont les valeurs propres de l'opérateur associé à l'observation. Ces valeurs propres sont les seuls résultats possibles. On suppose de plus que chaque état du système peut être représenté par une combinaison linéaire des états propres. Ces états propres forment une base de l'espace de Hilbert.

Enfin si l'on veut déterminer la probabilité du résultat d'une mesure, cette probabilité sera égale au carré du module du produit scalaire de cette valeur propre par le vecteur d'état, ou encore au carré du module de l'amplitude de probabilité associée au vecteur d'état (cette amplitude est égale au produit scalaire des valeurs propres par le vecteur d'état).

Il y a différentes manières d'interpréter ces probabilités. On pourrait dire que ces probabilités représentent la connaissance que nous avons du système. Ce serait une interprétation classique ou subjective de la probabilité. Cela montrerait que notre connaissance du système n'est pas complète. C'est là le sens du célèbre argument exposé par Einstein, Podolsky, et Rosen en 1935.

Partant du postulat réaliste suivant : « lorsque sans perturber en quoi que ce soit un système, nous pouvons prédire avec certitude (c'est-à-dire avec une probabilité égale à l'unité) la valeur d'une quantité physique, alors il existe un élément de réalité correspondant à cette quantité[5] », les auteurs veulent montrer que la MQ n'est pas complète.

Une théorie physique sera complète, si tout élément de la réalité physique a une contrepartie dans la théorie. Ils considèrent alors un système de deux particules qui ont interagi jusqu'à un certain temps t, après lequel elles se sont séparées sans plus interagir.

La MQ décrit le système des deux particules par une fonction d'onde unique. Supposons que l'on fasse une mesure sur l'une des particules après l'interaction.

[5] Physical Review, t.47,1935, p.777.

Que l'on mesure par exemple la quantité de mouvement de la particule 1. Les lois de la mécanique quantique nous donneront alors la quantité de mouvement de la particule 2. Sans donc avoir perturbé la particule nous avons pu prédire avec certitude la quantité de mouvement attachée à 2.

En vertu du postulat réaliste il y a un élément de réalité qui correspond à 2. Et réciproquement si nous mesurons la position de la particule 1 nous pourrons déduire la position de la particule 2. Or la MQ nous apprend que ces deux quantités sont des quantités conjuguées qui ne peuvent avoir de contrepartie simultanée dans la réalité. Il faut alors conclure que la MQ est incomp1ète.

> « Une description qui, comme celle qui est actuellement utilisée, est statistique en principe, ne peut être à mon avis que provisoire »[6].

Einstein oppose son réalisme à l'interprétation orthodoxe de la MQ. La réalité est indépendante de l'observateur, et la mesure introduit une perturbation qui rend la détermination des grandeurs conjuguées impossible. Néanmoins, en elles-mêmes les particules ont un moment et une position qui leur appartiennent en propre.

Je passe sous silence les développements du paradoxe EPR et sa reformulation par Bohm. Ce qu'il faut remarquer, c'est que ce dernier permet de mettre au point une expérience qui va permettre de trancher entre le réalisme (théories à variables cachées locales) et l'interprétation standard de la MQ (dite de Copenhague).

A l'heure actuelle, il semblerait que l'expérience donne raison à l'interprétation standard, et que la pensée devrait prendre en compte cette nouvelle "réalité", à savoir la non- séparabilité d'un système quantique.

§5. Plaçons-nous enfin du point de vue de la causalité, car la mécanique nouvelle semble avoir bouleversé nos idées sur cette notion. Rappelons que le principe de complémentarité nous permet de décrire le processus des interfé-

[6] Lettre d'Albert Einstein à Karl Popper, in *La logique de la découverte scientifique*, trad. Fr. p.471.

rences en utilisant l'image d'ondes. Mais si cette image permet de décrire, elle n'explique pas. Le nouveau formalisme permet de faire des prédictions, et "marche" très bien du point de vue des calculs. C'est un algorithme intéressant, mais on ne comprend pas le processus, on ne connaît pas la cause. Comme le dit Richard Feynman :

> « On peut se demander comment ça marche vraiment. Quel est le mécanisme en œuvre en réalité ? Personne ne connaît aucun mécanisme. Personne ne peut vous donner de ce phénomène une explication plus profonde que la mienne- c'est-à-dire une simple description »[7].
>
> L'indéterminisme devient fondamental en sciences physiques :
>
> « Le hasard est présent dans les lois fondamentales de la physique »
>
> (Feynman, ibidem).

§6. L'abandon du déterminisme et la victoire de l'interprétation standard de la MQ fait que l'on assiste à un triomphe du probabilisme. Mon but n'est pas de suivre la trace de ce probabiıisme, quelque intéressante que soit la question. Je ne vais dans ce mémoire m'intéresser qu'à la probabilité en elle-même, et à sa liaison avec le problème de l'induction.

Ma problématique est inspirée de la phrase un peu provocante de de Finetti : « la probabilité n'existe pas »[8].

L'auteur entend par là une entité ayant une existence séparée, objective, indépendante de nous. Bien plutôt, selon lui la probabilité est le degré de foi ou de croyance que nous attachons à nos jugements, propositions, prévisions. Les probabilités sont relatives à un individu et dépendent de l'état de nos connaissances. Elles peuvent être modifiées par des connaissances ultérieures. Bref elles n'appartiennent pas à la sphère de l'être, mais du sujet, de son opinion, de sa croyance.

---

[7] *La nature de la physique*, p.172.
[8] *Theory of probability*, préface, p. x.

Cependant ces probabilités, même en étant ce qu'il y a de plus arbitraire ou de plus subjectif dans le sujet, sont soumises à une seule condition : la cohérence, de t e l l e manière que les lois et les axiomes du calcul des probabilités leur soient applicables. Je me réserve de traiter la conception de de Finetti dans un chapitre à part.

A l'opposé de la conception de de Finetti se trouve la conception objectiviste représentée partiellement par la théorie fréquentiste de von Mises. Parmi les fréquentistes se range aussi Karl Popper, du moins celui de *Logik der Forschung*. Ce dernier exprime très bien la raison du choix d'une théorie objective de la probabilité. Ce choix correspond à une certaine conception de la science :

> « J'estime que seule une théorie objective peut rendre compte de l'application du calcul des probabilités à la science empirique »[9].

Or il est remarquable que tous les résultats et théorèmes du calcul des probabilités peuvent se déduire d'une théorie subjective de la probabilité. C'est ce qui est montré par de Finetti.

Mon étude sera en partie historique. Je commencerai par la notion de probabilité telle qu'elle est exposée chez Cicéron dans les *Académiques* (chapitre I). Le dialogue, on s'en souvient, est une exposition de la controverse qui oppose philosophes de la nouvelle Académie aux Stoïciens.

Je m'attacherai dans un deuxième temps à la naissance du calcul des probabilités (chapitre II) en montrant quelles notions les auteurs ont mises en œuvre pour quantifier l'incertitude et la traiter de manière mathématique. Je consacrerai aussi un chapitre au statut de la probabilité chez Laplace et l'usage heuristique qu'il en fait (chapitre III).

Je viendrai enfin au cœur du problème en m'attachant à la naissance de la conception fréquentiste chez Venn et à son développement chez von Mises (cha-

[9] *Logique de la découverte scientifique*, trad. Fr. §48, p.150.

pitre IV). J'exposerai en détail les principaux arguments qu'ils apportent à l'appui de 1eurs thèses.

Enfin un dernier chapitre est consacré à la conception personnaliste.

CHAPITRE I, PROBABILISME ET PROBABILITE : EMERGENCE DE LA PROBABILITE COMME THEME PHILOSOPHIQUE. LIEN AVEC LE SCEPTICISME.

§1. Le probabilisme est la doctrine qui affirme que sur le plan de la connaissance l'homme n'a pas accès à la vérité, qu'il doit seulement se contenter d'un savoir probable et que sur le plan de l'être il n'a accès qu'aux seuls phénomènes. Cette doctrine part du même constat que le scepticisme, à savoir l'impuissance de l'homme à atteindre le réel, et la précarité fondamentale de son savoir. Cependant elle se distingue du scepticisme. La thèse sceptique affirme en effet que l'homme n'a accès à aucune vérité quelle qu'elle soit :

« Il n'y a rien que l'homme puisse percevoir, rien qu'il puisse connaître, rien qu'il puisse savoir. Tout est enveloppé de ténèbres[10] ».

§2. La thèse sceptique est aussi ancienne que la philosophie, puisque nos principales sources, Cicéron[11] et Diogène Laërce, classent Parménide, Empédocle, Xénophane, Démocrite, Socrate, Platon parmi les précurseurs du scepticisme. Nous ne discuterons pas ici de la pertinence de ce classement, car c'est un point de critique historique, et nous renvoyons aux *Académiques* pour la discussion, ainsi qu'à Sextus Empiricus.

Le véritable fondateur du probabilisme comme doctrine philosophique fut Carnéade (vers 219-129 av. J.C.), philosophe de la Nouvelle Académie. Nous ne savons rien de ses doctrines, si ce n'est de manière indirecte par le témoignage de Cicéron, et de Sextus. Pour Carnéade, comme pour Arcésilas (vers 315-240 av. J.C.) qui l'a précédé de quatre

[10] Brochard, p.106

[11] *Académiques*, XXIII, 72 sqq.

générations à la tête de l'Académie, il n'y a pas de critérium de la vérité, quel qu'il soit : ni la raison, ni la sensation (αϊσθεσις), ni l'apparence (φαντασία) ne sont des critères de la vérité[12]. Il allait même douter de la certitude des propositions mathématiques. Cette critique radicale ruine la prétention stoïcienne à fonder la connaissance sur la représentation compréhensive (φαντασία καταληπτική).

En effet toute la théorie stoïcienne de la connaissance repose sur l'existence d'une telle représentation. Ce que nous traduisons par représentation, φαντασία, est aussi l'apparence ou l'image produite en nous par un objet sensible. Cette image est comparable à l'empreinte d'un cachet sur de la cire (Cléanthe[13]) ou à une altération de l'Âme (Chrysippe[14]) ; cette altération de l'Âme étant elle-même comparable à celle que subit le milieu extérieur qui transmet un son ou une couleur[15].

On pourrait penser à ce point de la théorie que le stoïcisme s'enferme dans une sorte de subjectivisme à la manière des Cyrénaïques, pour qui la réalité réside dans nos sensations de douleur et de peine. Qu'est-ce qui alors peut garantir la conformité de la φαντασία, *phantasia*, avec l'objet réel ? Qu'est-ce qui me dit que cette φαντασία est vraie ?

La représentation compréhensive va assurer ce rôle de critère de la vérité. Mais comment la reconnaître ? Les stoïciens soutiennent qu'elle possède des caractères intrinsèques qui la distinguent de toutes les autres, comme les serpents à cornes se distinguent des

[12] *Académiques*, XXIII, 72 sqq.
[13] Sextus Empiricus, *Contre les logiciens*, 1, 159.
[14] Diogène Laërce, VII, §6.
[15] Sextus Empiricus, op.cit., I, 372.

autres serpents[16]. Ces caractères ont une telle clarté, une telle évidence en eux-mêmes qu'ils ne peuvent produire qu'un assentiment invincible.

La représentation est compréhensive parce qu'elle entraine l'assentiment (συγκατάθεσις, *syncatathesis*).

Le texte suivant de Sextus en donne une définition (*Contre les logiciens*, 1, 248) : « Une représentation compréhensive est causée par un objet existant, et est imprimée et gravée dans l'âme de telle manière qu'elle ne puisse pas venir d'un objet non-existant[17] ».

Une telle représentation est une image fidèle de l'objet qui rend impossible toute confusion. « Une apparence plus claire et plus précise, voilà à quoi revient le critère[18] ».

Or c'est sur ce point précis que porte la critique de Carnéade, sur la non-distinction entre représentations compréhensives et représentations non-compréhensives (ακαταλεπτικαι, *acataleptikai*). Si aucune distinction n'existe, on sera alors obligé de conclure que rien ne peut être perçu.

La réfutation est construite sur un sorite : 1) il y a des représentations fausses, ce que tout le monde accorde, excepté les Epicuriens ; 2) mais l'objet d'une représentation fausse ne peut être perçu. Il est acata1epton ; 3) de plus, une représentation vraie, tout en restant la même, peut paraître fausse ; 4) enfin, parmi des représentations qui

---

[16] Sextus, idem, I, 227.
[17] Sextus, idem, I, 252.
[18] Cf. aussi *Académiques*, II, vi, 18 et xxiv, 77 :
« Ce qui est compris est une représentation (c'est le mot que nous avons usé hier pour rendre phantasia) imprimée et formée d'après l'objet d'où elle provenait, telle qu'elle ne pourrait être si cet objet n'existait pas. » (trad. Emile Bréhier, in *Pléiade*, p.196).

ne diffèrent en rien, il est impossible que les objets des unes soient perçus et les autres non. Donc il n'y a pas de représentation compréhensive, dont l'objet puisse être perçu[19].

Deux séries d'exemples viennent illustrer la thèse des Académiciens. Les premiers montrent que les fous, les ivrognes et les dormeurs ont des visions qui ne diffèrent en rien de celles des gens éveillés, sains d'esprit ou bien sobres, lorsqu'ils délirent. Les seconds montrent qu'il existe des objets indiscernables parce que trop semblables, comme les cent statues de bronze d'Alexandre, ou les cents empreintes de ce cachet imprimées dans la cire[20].

Une série d'objection peuvent être soulevées contre cette doctrine, mais la plus pressante semble être celle-ci : si rien ne peut être perçu, et si on ne peut rien savoir de manière certaine, et qu'il faille pratiquer la suspension du jugement, cela ne jette-t-il pas l'homme dans une inactivité coupable ?

Non, répondent les Académiciens par la voix de Carnéade, car si le Sage ne doit pas donner son assentiment à des représentations incertaines, faute de quoi il cesserait d'être sage, il pourra néanmoins avoir des opinions et agir en fonction de la probabilité ou de la crédibilité de ces opinions.

Ici se marque l'originalité de Carnéade. La conclusion de son scepticisme aurait dû être la suspension de jugement. Mais comme il faut agir dans la pratique, il faut un critère ou un guide pour la vie, une règle de vie (ratio vitae). Cette règle sera la représentation probable (φαντασια πιθανη, *phantasia pithané*).

---

[19] Emile Bréhier, *Chrysippe et l'ancien Stoïcisme*, p.97.
[20] Cicéron, *Académiques*, 11, xxvi, 63. Nous suivons en outre l'exposé de Victor Brochard, *Les sceptiques grecs*, Paris, 1887, réimprimé par Vrin, nouv. éd., 1986.

Cette représentation probable diffère de la représentation compréhensive parce qu'elle ne nous renseigne en rien sur la réalité de son objet : celui peut être existant comme non-existant. C'est pourquoi cette représentation probable pourra être fausse. C'est-à-dire ne correspondre à aucun objet existant.

Cette probabilité admet des degrés. Elle augmente lorsqu'elle n'est pas contredite par une autre représentation (aperispastos). Son degré le plus élevé est atteint lorsqu'elle a été examinée et éprouvée[21].

Quel est le caractère de cette probabilité ? Notons que c'est par ce mot, *probabilis* que Cicéron traduit le grec πιθανόν, *pithanon*. Cet adjectif est de la même famille que le verbe moyen-passif, πειθομαι, *peithomai*. Je suis persuadé, j'ai confiance, j'obéis. Il semble que Carnéade ait introduit ce terme pour éviter le terme technique assentiment (συγκατάθεσις, *syncatathesis*), qui avait une connotation trop stoïcienne. L'assentiment, nous l'avons vu, est l'impression invincible produite en nous par l'évidence d'une représentation claire et distincte. Cet assentiment précède la κατάληπσις, *catalepsis*, la prise ou la saisie par l'esprit de la chose. Il en est la condition nécessaire. Or soutenir qu'aucune chose ne peut être saisie ou perçue, ou comprise, c'est rejeter cet assentiment. Ce qui le remplace c'est la probabilité ou l'approbation (probatio) que nous donnons à la représentation.

On comprend que pour Lucullus ou Antiochus cette doctrine de la probabilité puisse paraître inconsistante. Elle contredit l'idée du sage stoïcien qui ne doit pas avoir d'opinions[22]. Ce dernier ne donne son assentiment qu'à des choses certaines et vraies.

---

[21] *Académiques*, II, xxviii, 91.

[22] Sextus Empiricus, *Contre les logiciens*, i, 166-184.

L'objection principale produite par Lucullus contre cette théorie de la probabilité est la suivante : si Carnéade nie qu'il n'y ait aucun critère de la vérité, (et il le nie), comment peut-il affirmer qu'une représentation est probable, ou proche de la vérité[23] ?

La réponse de Cicéron consiste à montrer que probable ne veut pas dire proche de la vérité, car des représentations probables peuvent être fausses.

Ainsi est introduite pour la première fois dans la philosophie, comme thème de réflexion la notion de probable. Carnéade lui donne en fait ses lettres de noblesse, car on se souvient que dans le *Gorgias* (501a), Socrate emploie πιθανὸν dans un sens négatif : qui se laisse facilement persuader, crédule.

Carnéade, substitue donc à la notion de vérité, qui est un absolu, la notion de certitude. La notion introduite est éminemment subjective, puisqu'elle est le degré de croyance que nous donnons à la représentation. Mais subjectif, ici comme ailleurs, ne signifie pas arbitraire. Une représentation n'est pas probable que parce que nous le croyons (c'est le premier degré de la probabilité, et peut-être le moins sûr), mais aussi parce qu'elle a été éprouvée et examinée. Notons enfin que ce n'est pas l'objet en lui-même qui est probable, mais c'est la représentation que nous en avons.

§3. Que devient le probabilisme après Carnéade ? La doctrine est liée au destin de l'Académie. Elle est continuée par Clitomaque, disciple de Carnéade, à qui nous devons l'exposition de la doctrine de son maître, via Cicéron. Puis, après la disparition de ce dernier, l'Académie semble verser dans l'éclectisme avec Philon de Larisse et

[23] *Académiques*, II, xviii, 60.

Antiochus d'Ascalon. Après ce dernier l'Académie semble disparaître. Il faut dire aussi qu'au premier siècle de notre ère le pyrrhonisme renait sous la direction d'Ænésidème, et que ce nouveau scepticisme va combattre le scepticisme académique.

CHAPITRE II, CHANCES ET PROBABILITES.

§1. Il serait fastidieux de suivre la trace de la notion de probabilité tout au long de l'histoire jusqu'au milieu du XVIIe siècle. Notons que le probabi1isme, comme théorie de la connaissance, connaît une longue éclipse jusqu'à sa résurgence au milieu du XVIIIe siècle. Cette éclipse peut s'expliquer par le triomphe du Christianisme.

En effet, n'oublions pas que le probabilisme partage des prémisses communes avec le scepticisme. Or la doctrine chrétienne est la négation de tout scepticisme. Ne suppose-t-e11e pas en effet un ensemble de croyances révélées au-dessus de tout doute ? Aussi voit elle de manière hostile l'émergence du scepticisme au XVI^{e} siècle.

Cependant on aurait tort de croire que l'Eglise ait totalement exclu le probable du champ de sa réflexion. La notion s'est en quelque sorte déplacée de la théorie de la connaissance vers la morale, la casuistique. Comme le dit Deman dans l'article probabilisme du *Dictionnaire de Théologie*, si les théologiens du Moyen-Age ne se sont pas contentés de l'opinion probable sur le plan de la connaissance scientifique, ils l'ont néanmoins accueillie dans le champ de la morale, comme directrice immédiate de l'action. Il y a chez eux une véritable science morale. Je ne retracerai pas ici le débat passionnant et passionné qu'a éveillé le probabilisme entendu comme doctrine morale. Tout ce dont je me contente de remarquer c'est le sens reçu du mot probabilité dans ce débat. Or ce sens est le sens cicéronien d'opinions approuvées par des gens doctes et sages.

Il n'y a rien de nouveau par rapport à ce que nous savons déjà. Bien plus intéressant est l'émergence d'un sens nouveau en rapport avec le calcul des chances naissant. C'est à cette liaison que nous allons consacrer ce chapitre.

§2. Les historiens se sont généralement accordés pour faire naître le calcul des probabilités dans le courant de l'année 1654 au moment où Pascal et Fermat résolvent indépendamment l'un de l'autre le problème de la « division des partis ».

Cependant, cette origine a été contestée par des historiens anglais. Leurs arguments sont les suivants : dès le XVIe siècle, on sait énumérer les différentes combinaisons de trois dés, et même de plusieurs dés (Tartaglia, *Trattato generale*, 1556), et on sait énumérer les différentes combinaisons qui donnent un point particulier. De là à l'idée de proportionner les mises aux chances de gagner un point, il n'y avait qu'un pas qui est franchi par le mathématicien Cardan.

M.G Kendall et F.N.David ont particulièrement développé ce point, et leur argumentation semble convaincante.
En effet, si l'on se tourne vers Cardan, on voit qu'il sait énumérer les différentes combinaisons de deux dés, de trois dés, qu'il sait énumérer les combinaisons qui amènent un point particulier, et qu'il recommande donc aux joueurs de proportionner leur mise à leurs "chances" respectives[24]. De plus il résout ce problème simple (un cas particulier du problème des dés proposé à Pascal par le Chevalier de Méré) : en combien de coups puis-je avoir une chance égale, ou puis-je parier 1 contre 1, qu'un point particulier va sortir ? Cependant sa méthode de calcul n'est qu'approximative. Elle repose sur un raisonnement qu'Ore

a appelé raisonnement sur la moyenne[25].De plus on peut relever chez lui d'autres erreurs significatives. A la fin de son livre il applique les raisonnements qu'il avait tenus sur les dés aux astragales et calcule le poids respectif de leurs faces en les supposant également probables ce qui constitue une induction fausse. C'est aussi prêter beaucoup trop à Cardan que faire de lui le précurseur de la loi des grands nombres comme le font Ore et Maistrov[26].

En résumé :

> « Le pas que lui ou Ferrari a franchi, et c'est un très grand pas, c'est d'introduire l'idée des combinaisons pour énumérer tous les éléments de l'ensemble fondamental des probabilités, et de noter que si tous les éléments de cet ensemble ont le même poids, alors le rapport du nombre des cas favorables au nombre total des cas donne un résultat en accord avec l'expérience[27].

S3. A la différence de Cardan, Galilée semble avoir un sens des concepts probabilistes digne de son génie[28]. En peu de mots en effet, il résout une apparente contradiction entre deux faits. Avec trois dés, les joueurs peuvent réaliser les points 9 et 12 en autant de manières que les points 10 et 11. Or ils savent qu'à la longue 10 et 11 sont plus avantageux que 9 et 12.
La réponse de Galilée est d'une simplicité et d'une limpidité remarquables :

> « Certains nombres sont plus facilement et plus fréquemment produits que d'autres, ce qui dépend de leur capacité à être formés d'un plus grand nombre de nombres »[29]

11 faut compter les combinaisons ordonnées et les considérer comme également possibles ("faciles" dit Galilée), plutôt

que de compter les combinaisons non ordonnées. Ce passage de l'un à l'autre n'était pas facile intellectuellement, comme l'illustre Hacking (op. cit. p. 51-52) en mettant l'exemple à jour. Il y a entre les deux la même différence qui existe entre la distribution de Maxwell-Boltzmann et la distribution Bose-Einstein en statistique.[30]

Ce sont les permutations et non les partitions qui sont également possibles. Les 6 partitions de 9 et 12 se résolvent en 25 permutations alors que les 6 partitions de 10 et 11 se résolvent en 27 permutations. L'avantage est de 27:25, soit 1/106. Celui qui a posé la question â Galilée, avait assez de connaissance empirique pour remarquer une différence de probabilité aussi faible.

Ce point a été discuté par Hacking et Maistrov. Ce dernier pense que l'allégation de Galilée selon laquelle les joueurs auraient détecté cette différence est une concession à une vieille tradition[31].

§4. Cette brève esquisse historique nous permet de conclure. Au XVIe siècle, on sait énumérer les différentes combinaisons possibles avec des dés. Ce sont là les prémisses du futur calcul des probabilités.
Le pas suivant ne sera réalisé qu'un siècle plus tard, lorsque Pascal associa les coefficients du binôme à la résolution du problème de la division des partis.

Ce problème, fameux entre tous, n'était pas nouveau au milieu du XVIIe siècle, puisque on en trouve la première mention dans la *Summa* de Fra Luca Pacioli, publiée en 1494, à Venise, et que Pacioli lui-même allègue une tradition antérieure[33].

Cependant il avait résisté aux efforts d'une génération de mathématiciens, parmi lesquels Cardan, Tartaglia, Peverone.

D'une manière simple on peut le formuler ainsi : deux Joueurs A et B jouent à un jeu quelconque en n parties. Pour une raison indéterminée, ils décident d'interrompre leur jeu avant la fin. Comment devront-être partagées leurs mises ? Le problème peut se généraliser à plusieurs joueurs.

La solution consiste à partager les mises selon les points qui restent à gagner, et suivant la probabilité que chaque joueur a de gagner.

Dans un premier temps Pascal suppose que les chances de chaque joueur sont égales. De Moivre généralisera par la suite le problème en tenant compte des inégalités de chance ou d'habilité entre les joueurs.

55. Quelle notion de probabilités utilisent les mathématiciens ? Sur quelle notion calculent-ils ?

Nous remarquons que jusqu'à présent, en esquissant cette brève histoire du calcul des probabilités avant 1700, nous avons fait peu usage du mot probable, ou du mot probabilité, ce qui semble paradoxal. Cela l'est moins, quand on connaît l'origine et l'histoire du mot probable que nous avons rappelées dans l'introduction de ce chapitre.

Les mathématiciens calculent sur des chances et non sur des probabilités. La différence est essentielle, et n'avait pas échappée à Cournot. La chance, du bas latin *cadentia*, dérivé du verbe *cadere* représente d'abord le point qui sort lorsqu'on jette les dés. C'est un événement qui appartient objectivement au dé de par sa nature propre, de par sa constitution physique.

On peut les énumérer, on peut les compter. Si le mot devient, par un glissement de sens progressif, synonyme du Sort, de la Fortune qui est allouée à chacun, il ne faut guère s'en étonner, étant donné l'usage mantique des dés[33].

La notion primitive dans le calcul des probabilités naissant est donc la notion de chance. Mais n'est-ce-pas plutôt la notion d'espérance mathématique ? Or la notion d'espérance n'est-elle pas subjective comme le souligne René Poirier ? « Un fait est probable lorsqu'il est attendu, prévu. Une assertion de probabilité fait en principe allusion à une expérience future. Or prévision, attente sont des faits capricieux, liés à mille causes : à l'état de notre corps, à notre imagination, à notre ennui »[34]

Si en peut argumenter que la notion d'espérance mathématique est une notion primitive du calcul des probabilités, comme le fait finement Lorraine Daston dans *The Empire of Chance*, on peut facilement soutenir que cette notion chez les mathématiciens est une notion subjective. En effet, dans le commentaire qu'il fait du livre de Huygens, Jacques Bernoulli précise : « le mot attente ne doit pas se prendre ici dans le sens ordinaire, selon lequel attendre ou espérer se rapporte à l'événement le plus favorable, quoique le contraire puisse arriver ; mais on doit entendre par ce mot l'espérance que nous avons d'obtenir le meilleur, tempérée et diminuée par la crainte du pire »[35]. Mais en fait comme la notion d'espérance mathématique définit les conditions d'un jeu juste, quelques soient les joueurs, on peut dire que cette subjectivité s'évanouit dans une intersubjectivité qui fonde le caractère objectif de la notion. Comme de

plus, cette espérance est proportionnelle à la chance de l'événement, conçue de manière objective, elle peut aussi prendre ce caractère objectif. Quoiqu'il en soit, cette difficulté à tracer de manière trancher une ligne entre les sens subjectifs et objectifs prouve l'ambiguïté du concept. Est-ce-que les auteurs du XVIIIe siècle échappent à cette ambiguïté ? Nous allons étudier la question chez le cas particulier de Laplace.

CHAPITRE III, PROBABILITES SUBJECTIVES ET OBJECTIVES CHEZ LAPLACE.

§1- L'œuvre de Laplace étant difficile pour celui qui n'a pas reçu une formation poussée en mathématiques, je ne traiterai que d'un aspect particulier de cette œuvre, à savoir la statistique mathématique, en relation avec le théorème de Bayes, l'usage des probabilités initiales uniformes, et les tests de signification.

Je traiterai aussi de l'interprétation de la probabilité chez Laplace, en essayant de voir si son déterminisme l'a conduit à une notion particulière de la probabilité. Pour une revue plus générale de Laplace, on pourra consulter Todhunter, pp.464-613, et les différents articles de Sheynin ainsi que Stigler, et Molina.

Laplace s'est intéressé dès le début de sa carrière scientifique à la notion de probabilité. Le premier mémoire qu'il lui consacre est présenté à l'Académie Royale des Sciences le 5 février 1772 (« Mémoire sur les suites récurro-récurrentes et sur leurs usages dans la théorie des hasards », 1774, réimprimé dans l'édition de ses Œuvres complètes, tome VIII, 5-24).

Bernard Bru présume que, s'il s'est intéressé si tôt - il n'a que 24 ans - au calcul des probabilités, alors que sa réputation scientifique était déjà établie, c'est parce qu'il trouvait dans la théorie des probabilités naissante un large domaine d'application de sa méthode d'intégration des équations aux différences finies partielles[36].

Ce mémoire est suivi d'un deuxième qui étend la matière de ce premier, et qui est lu à l'Académie les 10 et 17 mars 1773[37], où au milieu de développements hautement mathématiques, on trouve une dizaine de pages[38] dans lesquelles il définit les notions de hasard, de probabilité, d'espérance mathématique et morale. On y apprend ainsi que « nous regardons une chose comme l'effet du hasard, lorsqu'elle n'offre à nos yeux rien de régulier, ou qui annonce un dessein, et que nous ignorons d'ailleurs les causes qui l'ont produite[39]. » Dans cette définition nous retrouvons la constance des traits que Laplace prête au hasard, et qui lui permettra plus tard dans les applications du calcul des probabilités à la vie civile, de voir dans la supériorité du nombre de naissances des garçons sur celui des filles, l'effet d'une cause particulière, et non l'effet du hasard. On y trouve aussi dans ce mémoire une première expression du déterminisme :

« L'état présent du système de la Nature est évidemment une suite de ce qu'il était au moment précédent, et, si nous concevons une intelligence qui, pour un instant donné embrasse tous les rapports des êtres de cet Univers, elle pourra déterminer pour un temps quelconque pris dans le passé ou dans l'avenir la position respective, les mouvements, et généralement les affections de tous ces êtres[40]. »

Ces thèmes de réflexion philosophique seront repris dans le troisième mémoire de cette période que Laplace consacre au calcul des probabilités, le Mémoire sur la probabilité des causes par les événements. (O.C. t. VIII, 27-65).

Je vais limiter mon étude sur Laplace à ce dernier mémoire, au mémoire de 1778 publié en 1781, Mémoire sur les probabilités,

et aux chapitres III et VI du livre II de la *Théorie analytique des probabilités*.

Le deuxième mémoire de 1774 est important puisque, dès les premières pages Laplace y pose un principe dans lequel nous reconnaissons le théorème de Bayes. Ce théorème est important pour comprendre l'approche probabiliste et statistique de Laplace, c'est pourquoi j'y consacrerai quelques pages.

Remarquons que ce théorème est dû à Laplace. En effet, on ne trouve chez Bayes qu'une définition de la probabilité conditionnelle (propositions 3 et 5), et le problème suivant : trouver que la probabilité p d'un événement M se trouve dans un intervalle [a,b] sachant qu'un événement observé X est arrivé m fois sur n épreuves.

D'où Laplace a-t-il tiré ce principe ? A-t-il lu Bayes ? Tout ce que l'on peut dire à ce sujet se résume à des conjectures. Il est vraisemblable que Laplace soit parvenu de manière indépendante à ce résultat, en adoptant un raisonnement fiduciaire (Stigler), Bayes étant cité seulement en 1781 sur le continent dans une introduction anonyme à un mémoire de Laplace.

Le principe est posé au début du mémoire de 1774 sans démonstration. Laplace sera plus rigoureux dans la *Théorie analytique des probabilités* au chapitre I du livre II où le théorème de Bayes est déduit de la règle de multiplication des événements composés et de la définition de la probabilité conditionnelle qui en est le corollaire[41].

Laplace énonce ainsi le principe au début du Mémoire de 1774 :

« Si un événement peut être produit par un nombre n de causes différentes, les probabilités de l'existence de ces causes

prises de l'événement sont entre elles comme les probabilités de l'événement prises de ces causes, et la probabilité de l'existence de chacune d'elles est égale à la probabilité de l'événement prise de cette cause, divisée par la somme de toutes les probabilités de l'événement prises de chacune de ces causes[42] » Cet énoncé, quelque peu ésotérique pour un profane, se comprend mieux, si, d'une part, on reproduit la déduction de Laplace, et si, d'autre part, on l'illustre.

Laplace part de la règle de multiplication des probabilités, qui était connue de De Moivre, et peut-être de Cardan. Cette règle dit que lorsque deux événements sont indépendants, la probabilité de l'événement composé de ces deux événements est le produit de leurs probabilités respectives.

Par exemple dans un jeu de 32 cartes le tirage (avec remise) d'un roi est indépendant du tirage d'un as. Et la probabilité du tirage d'un roi et d'un as (événement composé de deux événements élémentaires) est égale à 1/16.

Mais que se passe-t-il lorsque la probabilité d'occurrence d'un événement dépend d'un autre déjà arrivé ? La règle du produit s'applique encore, mais avec quelques amendements.

La probabilité de l'événement composé sera alors égale au produit de la probabilité du premier événement par la probabilité du second, sachant que le premier est arrivé.

Symboliquement, P(AH)= P(A)xP(**H**|A). Cette règle de multiplication définit la probabilité conditionnelle de H étant donné A. La règle permet aussi d'évaluer la probabilité d'un événement futur.

En effet, il faut considérer comme événement composé un événement observé et un événement futur qu'on attend. Soit E un

événement passé (observé) dont on connaît la probabilité V, et soit e un événement futur dont on se propose de déterminer la probabilité P. Soit E+e un événement composé de E et de e, dont la probabilité est v. D'après la règle de multiplication v=VxP, d'où l'on tire P=v/V.

Le théorème de Bayes est déduit de la même façon. Il considère comme événement composé l'événement observé résultant d'une des causes. La probabilité de cet événement composé sera égale au produit de la probabilité de l'événement observé par la probabilité que cet événement ayant eu lieu, la cause dont il s'agit existe. On tire de là la probabilité de la cause tirée de l'événement. P(C|A), qui est une fraction égale à P(CA)/P(A) (CA événement composé). On exprime ensuite numérateur et dénominateur en fonction des causes C_i. P(CA) est alors égale à P(C)xP(A|C) (probabilité que cette cause ayant lieu, l'événement A arrivera), et P(A) est égale à $\sum_j$P(A|Cj)xP(Cj). D'où une expression familière pour un lecteur contemporain du théorème de Bayes.

L'énoncé de Laplace diffère cependant de l'énoncé moderne parce qu'il suppose égales les probabilités des différentes causes[41]. Mais il considère aussi le cas où toutes les causes ne sont pas a priori possibles. Il rencontre alors l'énoncé moderne[44].

Voyons l'usage que Laplace fait du théorème. Il l'applique au problème suivant, qui est d'estimer une proportion inconnue x de billets blancs et de billets noirs à partir de tirages répétés de billets (chaque fois avec remise). Ce problème est en substance identique à celui de Bayes. Sa résolution mène aux mêmes formules numériques. Exprimé symboliquement cela donne un

rapport de deux intégrales eulériennes. Le problème, non résolu par Bayes pour de grandes valeurs de n (nombre de répétitions) était d'évaluer l'intégrale du numérateur. Pour cela Laplace a besoin d'un théorème auxiliaire :

"On peut supposer les nombres p et q tellement grands, qu'il devienne aussi approchant que l'on voudra de la certitude que le rapport du nombre de billets blancs au nombre total des billets renfermés dans l'urne est compris entre deux limites (p/p+q)-w et (p/p+q)+w, w pouvant être moindre qu'aucune grandeur donnée"[45].

Le rapport empirique p/p+q est l'estimateur du rapport inconnu x. Laplace en termes modernes, affirme la consistance a posteriori du rapport p/p+q[46]. Mais comme l'espérance mathématique de la distribution a posteriori de x est égale à p+1/p+q+2, cet estimateur est seulement asymptotiquement sans biais[46].

Je ne refais pas la démonstration de Laplace (voir Stigler 196S, p.133). Elle repose sur sa méthode d'évaluation asymptotique des intégrales définies. Laplace donne de plus une formule qui permet d'évaluer la différence entre l'unité et la probabilité précédemment calculée. Pour l'instant Laplace demeure dans le domaine de la pure analyse, et on ne voit pas les applications pour les phénomènes naturels. Pour cela il faut des formules plus pratiques.

Le mémoire de 1781 va les fournir. Mais les formules les plus faciles à utiliser se trouveront dans la *Théorie analytique des probabilités*, quelques 40 ans plus tard (démonstration du théorème central limite). J'étudierai les formules du mémoire de 1781. puisque le raisonnement est semblable à celui du mémoire de 1774.

Le problème qui sert à introduire ces formules dans le mémoire de 1781 ressemble au problème des tickets dans une urne : « Supposons que l'on ait observé que. sur p+q enfants, il est né p garçons et q filles, et que l'on cherche la probabilité P que, sur m+n enfants qui doivent naître il y aura m garçons et n filles...»[48]

Ce type d'inférence est typique de la méthode de Laplace. Soit x la probabilité qu'un enfant qui va naître soit un garçon, et (1-x) la probabilité qu'il soit une fille. Ici il faut remarquer qu'il y a un flottement dans le vocabulaire de Laplace, car cet x inconnu pourrait aussi bien être la cause des événements observés (ici garçons et filles déjà nés), ou encore la possibilité de tels événements.

Je reviendrai sur ce point délicat lorsque je traiterai de l'interprétation des probabilités. La distribution a posteriori de x est égale par le théorème de Bayes à :

$$\frac{x^p(1-x)^q\,dx}{\int x^p(1-x)^q\,dx}$$

La probabilité de l'événement futur (naissances de m garçons et de n filles) s'obtient par les principes déjà exposés en changeant les exposants du numérateur en p+m et en q-r, respectivement, et en transformant celui-ci en intégrale. Nous avons alors un rapport d'intégrales qui est simplifié en pmqn/(p+q)m+n en utilisant la formule de Stirling. L'analyse est exactement la même dans les deux mémoires[49].

> « Cette valeur de P est la même que celle à laquelle on parviendrait en supposant les possibilités des naissances des garçons et des filles dans le rapport de p à q; d'où

> il est naturel de conclure que ces possibilités sont à très peu près dans le même rapport, et qu'ainsi la vraie possibilité de la naissance d'un garçon est très approchante de p/p+q […] On peut tellement faire croître p et q que la différence de P à l'unité soit moindre qu'aucune grandeur donnée […]. On voit par là comment les événements, en se multipliant, nous indiquent d'une manière de plus en plus probable leur possibilité respective »[50]
>
> On reconnaît dans ce paragraphe l'expression d'une loi des grands nombres que von Mises appelle seconde loi des grands nombres[51]. Cette loi des grands nombres chez Laplace apparaît comme un théorème mathématique, conséquence du comportement asymptotique d'intégrales, plutôt que comme un énoncé empirique concernant le comportement des choses du monde.

Laplace approche ensuite cette probabilité par une série convergente dont le premier terme donne la limite supérieure et la somme des deux premiers termes la limite inférieure de l'expression. Ainsi, on pourra toujours déterminer avec quel résidu d'incertitude on se sera approché du rapport cherché.

Comme ce théorème très important lui permet de calculer la probabilité que x est compris dans certaines limites, il lui donne aussi la probabilité que x soit inférieur à une limite. Cela lui permet de construire un test d'hypothèses unilatéral. Considérons le problème suivant de philosophie naturelle, qui est le plus de facilité des naissances des garçons par rapport à celles des filles. Cette différence est-elle due au hasard ou à une cause particulière ? Il faut montrer que les naissances des garçons sont plus possibles que les naissances des filles et ce

avec une probabilité écrasante. Laplace calcule la probabilité que la possibilité de la naissance d'un garçon soit inférieure ou égale à un demi. Cette probabilité est de l'ordre de 10^{-42}.

> « On peut regarder comme aussi certain qu'aucune autre vérité morale, que la différence observée à Paris entre les naissances des garçons et celles des filles est due à une plus grande possibilité dans la naissance des garçons. »[52]

Cependant, même si le mémoire de 1781 offre des formules qui permettent de calculer avec quelle probabilité une estimation se rapproche de la grandeur estimée, ces formules sont encore peu commodes pour l'usage. C'est pourquoi dans les trente ans qui suivent ce mémoire, Laplace s'attache à trouver d'autres formules asymptotiques. Une de ces formules, avec le théorème central limite, est d'une importance particulière dans le calcul des probabilités, et dans la statistique. Il s'agit du théorème de De Moivre-Laplace, qui dit qu'une distribution binomiale tend asymptotiquement vers une distribution normale. Ce théorème avait déjà été démontré par De Moivre pour les distributions binomiales symétriques de la forme (1+1)n.

Le problème, déjà étudié par Jacques Bernoulli, consiste à évaluer le rapport du plus grand terme du développement du binôme à la somme de tous les termes. Puis d'évaluer la probabilité d'un écart de « 1 » par rapport au terme central de la distribution. Cette probabilité est égale à l'unité (l'aire totale de la distribution), moins les deux aires qui représentent les deux queues de la distribution. En termes symboliques. P(X>t)=1-2N(t), où N(.) représente la loi normale centrée réduite. Bien entendu, la démonstration complète de Laplace est plus

difficile, et diffère quelque peu de cet énoncé simplifié (se reporter au chapitre III du *Traité analytique*), mais en substance les évaluations numériques conduisent aux mêmes résultats.

Par exemple Laplace montre que si le rapport du nombre de naissances des garçons à celui des filles est de 18 contre 17, alors sur 14000 naissances la probabilité que le nombre de garçons soit compris entre 7037 et 7363 est égale à 0,994303[53].

Bien que ce dernier exemple ne soit pas dans la ligne bayésienne, ces quelques exemples auront montré l'approche bayésienne de Laplace. Il y aurait à considérer d'autres domaines d'applications, notamment en astronomie où il faut déduire des observations une loi des erreurs. Cependant, je m'arrête ici, et me tourne vers l'interprétation philosophique des probabilités.

§2. Laplace est connu pour l'affirmation universelle du déterminisme. Des auteurs ont conclu que ce déterminisme a empêché de plus larges applications du calcul des probabilités aux sciences de la nature[54]. De plus dans le cadre de ce déterminisme, la probabilité ne peut prendre qu'un sens subjectif :

> « La courbe décrite par une simple molécule d'air ou de vapeurs [sic] est réglée d'une manière aussi certaine que les orbites planétaires : il n'y a de différence entre elles que celle qu'y met notre ignorance. La probabilité est relative en partie à cette ignorance, en partie à nos connaissances. »[55]

Cependant si les probabilités n'étaient que l'expression de notre ignorance, comment aurions-nous pu calculer sur elles ? Cette apparence de paradoxe se résout aisément[54]. Les probabilités ne sont qu'<u>en partie</u> dues à notre ignorance. Autrement

dit il y a en elles quelque chose de réel qui n'est pas que le reflet de notre ignorance.

La théorie classique reconnaissait déjà ce quelque chose de réel dans la définition qu'elle donnait de la probabilité : rapport du nombre de cas favorables à l'événement à l'ensemble de tous les cas possibles (favorables et défavorables). Cette définition fait intervenir la notion de cas, qui, comme nous l'avons vu, représente la combinaison qui va donner le point, et qui est conçue de manière objective.

Le problème nouveau, surgi avec Bernoulli, était non pas la détermination des cas possibles dans les jeux de hasard, détermination a priori aisée avec le calcul combinatoire, mais bien la détermination a posteriori des cas à partir de l'expérience, là où les cas possibles n'étaient pas énumérables directement, comme dans les calculs qui font intervenir les annuités, ou les calculs de l'espérance de vie.

Dans le cas des jeux de hasard, la notion de probabilité était de ce point de vue une notion objective, ce qui était attesté par les auteurs qui parlaient plutôt du calcul des chances que du calcul des probabilités.

Quant à Laplace il est l'héritier de ces premiers probabilistes que sont Jacques Bernoulli et De Moivre, mais il intègre aussi dans son analyse l'apport de Bayes, l'usage du théorème ainsi nommé, et l'usage de distributions initiales uniformes lorsque nous n'avons aucune donnée a priori sur la loi initiale, ce que Keynes appellera plus tard le "principe d'indifférence". Par conséquent, on peut dire qu'implicitement il reconnaît la différence qui existe entre une probabilité

objective et une probabilité subjective, bien qu'il n'emploie pas cette terminologie moderne qui lui est postérieure.

La première, il l'appelle possibilité ou facilité, et la seconde probabilité. La possibilité des événements est souvent inconnue. Néanmoins on peut la déterminer de trois manières : « [...] A priori, lorsque par la nature même des événements, on voit qu'ils sont possibles dans un rapport donné [ce cas s'applique principalement aux jeux de hasard] ; c'est ainsi qu'au jeu de croix ou pile, si la pièce qu'on jette en l'air est homogène et que ses deux faces sont entièrement semblables, on juge croix et pile également possibles : […] a posteriori, en répétant un grand nombre de fois l'événement dent il s'agit, et en examinant combien de fois il est arrivé ; […]enfin par la considération des motifs qui peuvent nous déterminer à nous prononcer sur l'existence de cet événement[…].Le premier de ces moyens donne la possibilité absolue des événements ;le second la fait connaître à peu près, comme nous le ferons voir par la suite, et le troisième ne donne que leur possibilité relative à l'état de nos connaissances.»[57]

Laplace semble embarrassé par les distinctions qu'il vient de faire. Ne sont-elles pas artificielles ? Faut-il les prendre au sérieux dans le cadre du déterminisme ? C'est pourquoi, il précise :

> « Chaque événement étant déterminé en vertu des lois générales de cet univers, il n'est probable que relativement à nous, et par cette raison, la distinction de sa possibilité absolue et de sa possibilité relative peut paraître imaginaire… »[58]

> Cette distinction n'est cependant pas imaginaire, si on prend soin de distinguer les circonstances variables qui concourent à la production des événements, comme le mouvement que la main imprime aux dés, des circonstances constantes, comme l'habileté des joueurs ou la pente des dés à retomber sur une des faces plutôt que les autres.
>
> Les premières, Laplace les appelle hasard. Il y aurait place dans cette brève indication pour une notion de hasard objectif chez Laplace, conçu non plus comme une absence de causalité, ou de finalité, mais comme la réunion d'une multitude de causes particulières.

Je n'exploiterai pas cependant cette indication, pour m'arrêter au second type de circonstances. Ces circonstances constantes forment la possibilité absolue des événements (ibidem), et leur connaissance plus ou moins étendue leur possibilité relative.

Il semblerait que Laplace entende la possibilité absolue dans un sens aristotélicien plutôt que dans un sens leibnizien, synonyme de non-contradictoire.

En effet est possible absolument un événement qui pourrait se réaliser[59], qui a la puissance de se réaliser, mais qui ne se réalise que si des causes viennent s'adjoindre à cette puissance pour la faire passer à l'acte. C'est ainsi que j'interprète le passage suivant de Laplace : « Seules [les circonstances constantes qui forment la possibilité absolue des événements], elles ne suffisent pas pour les produire : il est de plus nécessaire qu'elles soient jointes aux circonstances variables dont j'ai parlé ; elles ne font ainsi qu'augmenter la

probabilité des événements, sans déterminer nécessairement leur existence.

Le lien entre probabilité et possibilité demande à être éclairci. On pourrait dire que les probabilités et les possibilités ne diffèrent pas en nature, mais seulement en degré. Un événement sera possible s'il a un degré quelconque de probabilité, aussi petit soit-il. Et si ce degré devenait infiniment petit, l'événement deviendrait improbable, et non impossible. Néanmoins, bien des auteurs, en particulier Cournot et Borel, ont estimé que des événements avec une probabilité très faible étaient impossibles (principe de Cournot). Emile Borel considère quelles sont les probabilités négligeables à l'échelle humaine, puis à l'échelle cosmique en comparant ces probabilités aux dimensions de l'univers[61].

Cependant il existe une différence de nature, car les possibilités se rapportent aux choses, et les probabilités dépendent toujours d'un sujet, sans être toutefois complètement subjectives. Elles ont parfois un fondement objectif, lorsqu'elles dépendent de la possibilité objective des événements. Les probabilités entièrement subjectives chez Laplace résulteraient alors de l'absence totale de données sur les événements, lorsque l'ignorance est complète et que nous devons supposer les événements équiprobables.

Dans ses analyses et exemples, Laplace mêle souvent les deux points de vue subjectif et objectif. De plus le théorème de Bayes permet de concilier les deux : Le point de vue subjectif dans les probabilités a priori uniformes, et le point de vue objectif dans la distribution a posteriori qui a pris en compte

les données (« les événements », d'après le vocabulaire laplacien).

Les deux exemples suivants le montrent, quoique n'étant pas situés dans la ligne d'analyse bayésienne. Dans ces deux cas, l'objectivité est tirée de l'examen même des choses, comme l'homogénéité des faces d'un dé, ou le nombre de fois où un événement a été observé (sa fréquence). Et la subjectivité vient de l'ignorance où nous nous trouvons face aux causes des phénomènes. Nous devons alors les supposer également probables. Laplace nous avertit du danger qu'il y aurait dans le calcul de ne pas prendre en compte cette différence qui existe entre les deux probabilités.

Considérons deux joueurs A et B. dont les adresses respectives sont inconnues et qui jouent à un jeu quelconque, que nous supposerons être un tournoi de tennis. On se propose de calculer la probabilité que A gagnera les trois sets. Si on ne connaît rien sur A la probabilité qu'il a de gagner le premier set est 1/2. Et la probabilité qu'il a de gagner les trois sets est $1/2^3$. Ce calcul est fondé sur l'égalité supposée entre les deux joueurs. Cependant cette égalité n'est pas parfaite puisqu'il y a nécessairement un joueur plus habile, celui qui va gagner le tournoi. Laplace représente par (1+a)/2 la probabilité du joueur le plus fort de gagner la partie, et par (1-a)/2 celle du plus faible. Comme on ne sait pas si A est le plus fort ou le plus faible, sa probabilité de gagner les trois parties sera :
$1/2^3$ $[(1+a)^3+(1-a)^3]$. Cette probabilité sera plus grande que celle calculée précédemment.

Le second cas est tout aussi intéressant. Il s'agit du célèbre problème de St-Pétersbourg. Il faut calculer l'espé-

rance d'un joueur A qui reçoit un franc s'il amène pile au premier coup, deux francs s'il amène pile au second coup, …2" francs s'il amène pile au nième coup. La solution conduit àune équation aux différences finies. Laplace commente cette solution :

> « Nous supposerons dans cette solution que la pièce qui, jetée en l'air, doit amener croix ou pile n'a pas plus de pente pour amener l'un plutôt que l'autre ; or cette supposition n'est admissible que mathématiquement, car physiquement il doit y avoir une inégalité ; mais comme les deux joueurs A et E ignorent en commençant le jeu de quel côté est la plus grande pente, on pourrait croire que cette incertitude n'augmente point et ne diminue point leur avantage. On va voir cependant que rien n'est moins fondé que cette supposition. »[62]

Je passe les calculs numériques qui résultent « de la supposition que la pièce a une plus grande pente à tomber d'un côté plutôt que de l'autre ». Laplace conclut que le calcul de l'espérance d'un des joueurs serait impossible si on ne savait pas dans quelles limites était comprise la pente.

En résumant ce passage, je dirai qu'une inégalité de pente entre les faces d'un dé est la cause d'une fréquence d'apparition plus grande de cette face.

> « […] Souvent entre les faces d'un dé, qui semble parfaitement cube, il existe une inégalité de pente très sensible, en sorte que, sur un fort grand nombre de coups, une des faces arrive plus souvent que l'autre, ce qui vient et de l'hétérogénéité de la matière du dé et de ce que sa figure n'est pas exactement cube. »[63]

Laplace conclut : « Ces exemples suffisent pour faire voir avec quelle précaution on doit appliquer aux objets physiques les considérations mathématiques sur le Calcul des probabilités. On suppose dans la théorie que les différents cas qui amènent un événement sont également probables, ou, s'ils ne le sont pas que leur probabilité est dans un rapport donné. Quand on veut faire usage de cette théorie, on regarde deux événements comme également probables, lorsqu'on ne voit aucune raison qui rende l'un plus probable que l'autre, parce que, quand bien même qu'il y aurait une inégale possibilité entre eux, comme nous ignorons de quel côté est la plus grande pente, cette incertitude nous fait regarder l'un comme aussi probable que l'autre. »[64]

§3. Que conclure de notre côté de ce rapide examen ? D'une part, Laplace a donné sa prédilection à des distributions initiales uniformes. Mais, comme le remarque Stigler, ce n'est pas là le signe d'une soumission à une hypothèse métaphysique, mais seulement un artifice de calcul, comme le montre l'exemple de la distribution non-uniforme que nous avons examinée. D'autre part les probabilités ne sont pas que fondées sur notre ignorance. Elles peuvent être fondées sur des événements observés, ce qui, d'un point de vue positiviste leur donne leur objectivité. Mais Laplace n'est pas un positiviste. Par conséquent ses critères d'objectivité sont autres.

Il est intéressant de voir que l'époque qui suit Laplace est l'âge du positivisme. Comte veut fonder une science positive sur des faits ou sur ce qui est observable. D'autre part l'empirisme devient une tradition très vive en Angleterre. Ce

contexte philosophique et le discrédit qu'avaient jeté les travaux de Condorcet, Laplace et Poisson sur les décisions des cours de justice font que des auteurs font chercher un fondement plus ferme à la probabilité, non plus dans la supposition d'égalité des chances, et dans le principe d'indifférence qui la fonde, mais dans la fréquence avec laquelle un phénomène se répète. C'est le cas de Venn, que nous allons examiner dans le chapitre suivant.

CHAPITRE IV FREQUENCES ET PROBABILITES

§1. Dans une interprétation objectiviste, la probabilité est une caractéristique physique d'un événement, que l'on peut déterminer de plus en plus précisément par les fréquences d'apparition de l'événement dans les séries d'épreuves répétées. On définit la fréquence relative comme le rapport du nombre d'apparitions de l'événement sur le nombre total des épreuves.

Il faut s'assurer que ce ratio soit stable, que les fréquences convergent vers une valeur limite que l'on déterminera de manière approximative. A la longue, les fluctuations du hasard se stabilisent pour faire apparaître une régularité. Ceci est présenté comme un fait empirique. On retrouve ce leitmotiv chez tous les fréquentistes, à commencer par Venn :

> « La notion fondamentale que le lecteur doit fixer dans son esprit … est … la notion de série. Mais c'est une série d'un genre spécial …. Elle combine l'irrégularité des individus avec la régularité de l'ensemble … En dépit du désordre individuel, un ordre collectif commence à prévaloir. »[65]

> « La propriété de la stabilité des fréquences […] n'est pas la conséquence d'une déduction logique … Que les fréquences … possèdent effectivement cette propriété est un fait empirique ou observationnel basé littéralement sur des millions d'observations. »

Je vais m'attacher à l'origine de ces idées que l'on trouve de nos jours dans la littérature probabiliste en me tournant en premier lieu vers Venn.

§2. Venn est un auteur suffisamment intéressant pour qu'on s'y arrête. Son livre. *Logic of chance*, a connu trois éditions de 1666 à 1686. C'est une livre qui tranche par rapport aux autres traités de probabilité de l'époque, en ce sens que ce n'est pas un traité mathématique sur les probabilités, mais une exploration des fondements, la première du genre semble-t-il, de la théorie des probabilités. Le livre, dans sa troisième édition, est divisé en trois parties. La première porte le titre très significatif de "Physical foundations of the science of probability" (ch.1-5). La seconde s'intitule "Logical superstructure of the above physical foundations", (ch.6-14), •t la troisième concerne "Various applications of the theory of gambling" (ch.15-19), et semble indépendante du reste.

Le but général de l'auteur n'est pas de construire une axiomatique de la probabilité, conçue comme une branche particulière de la logique, mais d'explorer les fondements physiques de la théorie afin de mieux la fonder sur l'expérience.

En effet la théorie précédente, que l'on peut appeler théorie classique pour faire court, et dont le représentant le plus éminent est Laplace, souffre à ses yeux de plusieurs défauts. Le premier est la définition même de la probabilité, conçue comme rapport des cas

favorables à l'événement sur l'ensemble des cas possibles. Cette définition ne met pas assez l'accent selon lui sur le fondement physique de la probabilité, en ce sens qu'elle est fondée sur le *principe de raison insuffisante* (que Keynes appellera le *principe d'indifférence*), qui seul permet de transformer a priori les cas possibles en cas également probables. Le second défaut de la théorie classique, et celui- là n'est pas imputable à Laplace, vient de l'usage qu'a fait Quetelet de la loi normale, comme on l'appelle aujourd'hui, privilégiant ainsi de manière arbitraire une famille de distributions, celles qui sont unimodales et symétriques par rapport au terme centrai (chapitre 2).

Venn, comme il le dit lui-même est inspiré par un point de vue matériel en logique (préface de la première édition, p.x), qu'il oppose au point de vue formel ou conceptualiste représenté par De Morgan. Il critique ce dernier parce qu'il a un point de vue conceptuel sur la probabilité, et parce qu'il introduit trop d'éléments subjectifs dans la science. C'est donc inspiré par ce point de vue "matériel" qu'il met l'accent sur la notion de série.

Cette série est un agrégat d'individus, ou de choses, ou d'attributs, ou de substances, ou d'événements, Venn semblant peu soucieux d'entrer dans des querelles métaphysiques sur la définition de mots qui présente deux caractères particuliers que nous avons déjà signalés, à savoir le désordre individuel et l'ordre collectif.

Le désordre individuel engendre l'incertitude dans la prévision de l'occurrence d'une instance (p.**5**). Par exemple la durée de vie de cet individu-ci est inconnue, donc incertaine,

mais en revanche, la durée moyenne de la vie de ce groupe d'hommes varie peu. Cette moyenne est dotée d'une fixité -relative- que n'ont pas les instances individuelles.

Toutefois cette moyenne n'a pas chez Venn la fixité d'un type. Il critique Quetelet sur ce dernier point. Elle est soumise à des fluctuations, qui à la longue, détruisent l'uniformité observée. Cette variabilité de la Nature, due à des causes cachées et multiples (chapitre 3), est un obstacle qui empêche de prendre des séries trop grandes, dans lesquelles nous ne retrouverions plus l'ordre recherché, mais le désordre. Seuls les jeux de hasard offrent une situation idéale dans laquelle les fluctuations tendent vers une limite qui ne varie pas. Cette notion de limite ne prend sens que lorsque nous substituons une série idéale à une série réelle.

Dans sa critique de la théorie subjective de la probabilité, représentée par De Morgan à cette époque, Venn est plutôt nuancé. Il ne nie pas la probabilité définie comme degré de croyance que nous attachons à des propositions ou à des événements, au profit exclusif d'une probabilité définie comme une fréquence statistique, mais il pense que la première est soumise à de telles variations qu'il est tout à fait impossible qu'elle se prête à une évaluation quantitative. Détachée de l'expérience, la science des probabilités deviendrait une branche de la psychologie.

Pour fonder une science générale de l'inférence, la notion de probabilité doit être objective. La probabilité subjective devient alors un simple appendice de la probabilité objective (p.138). En résumé une probabilité subjective non rapportée à

une fréquence statistique perd de sa justification. Comme le remarque Jean Largeault :

« L'interprétation fréquentiste, qui identifie probabilité et fréquence relative […] semble être suggérée par une attitude vérificationiste en épistémologie. Cette attitude vérificationiste se retrouve naturellement chez l'autre éminent représentant de la tendance fréquentiste, von Mises.

§3. Richard von Mises a consacré son œuvre scientifique à la mécanique et au calcul des probabilités. Il est le continuateur de Venn dont il dit s'inspirer (p.22).

L'œuvre de von Mises peut être divisée en deux parties à des fins d'exposition. Une partie est critique et l'autre est constructive.

La partie critique porte sur la théorie classique et sur la définition de la probabilité comme rapport de tous les cas favorables à l'ensemble des cas possibles. « Les auteurs, dit-il, commencent avec les cas également probables pour seulement abandonner ce point de vue à un moment propice et se tourner vers la notion de probabilité fondée sur la définition fréquentiste ; en fait ils introduisent parfois explicitement une telle définition.

La définition classique n'est pas circulaire pour von Mises, mais elle réduit toutes les distributions de probabilité à une distribution uniforme et c'est là son principal défaut. Or les distributions uniformes n'existent pas toujours (p.69). Prenons le cas d'un dé biaisé. Le plus léger éclat à l'un des coins pourrait briser cette distribution initiale. On ne peut

plus par conséquent appliquer les règles du calcul (addition et multiplication).

Il existe une autre série de problèmes non résolus par la théorie classique. Un de ces problèmes est la probabilité de la mort. Une table d'une compagnie d'assurance donne pour la probabilité de mourir d'un homme de 40 ans dans l'année 0,011. Où sont les cas "favorables" dans cet exemple ? Est-ce qu'il y a 1000 possibilités parmi lesquelles 11 sont favorables, ou bien y a-t-il 3000 cas possibles parmi lesquels 33 sont favorables ? Il est inutile de chercher dans les manuels une réponse à ces questions (p.70).

Dans tous les cas comme ceux-là, « les auteurs passent, comme si cela n'avait aucune importance, de la considération de probabilités a priori à la discussion de cas où la probabilité n'est pas connue a priori, mais doit être trouvée a posteriori en déterminant la fréquence de différents attributs dans une longue série d'expériences. »[69]. Avec une extraordinaire intrépidité, ils assurent que les théorèmes valables pour les probabilités a priori sont aussi valables pour les probabilités a posteriori. Et si certains veulent justifier cette étape, ils se réfèrent généralement à la loi des grands nombres qui est supposée fournir un pont entre les probabilités a priori et celles qui sont déterminées par l'expérience (ibidem). Donc tôt ou tard les auteurs classiques ne peuvent s'empêcher d'introduire, parallèlement à la définition classique, un concept de probabilité fondé sur des fréquences relatives. Il n'y a pas de raison par conséquent de donner son adhésion à une définition qui est trop étroite pour inclure la plupart des applications de la statistique (p.70).

Ces critiques sont similaires en substance à celles de Venn. De même lorsqu'il examine le côté subjectif de la probabilité, il reconnaît qu'il ne serait pas impossible de mener une investigation psychologique détaillée des fondements subjectifs de nos estimations de probabilité, mais leur rapport au calcul des probabilités serait semblable à celui de nos sentiments de température à la thermodynamique scientifique. »[70] Les observations répétées et les déterminations de fréquences sont les thermomètres de la théorie des probabilités (ibidem).

Il y a un point qui n'est toutefois pas traité dans le détail par Venn et sur lequel von Mises attire notre attention. Selon les auteurs classiques, la loi des grands nombres fournirait un pont entre les probabilités calculées a priori et les probabilités fondées sur des fréquences. Ceci est énoncé de plusieurs façons dans la littérature. Venn le rapporte ainsi : « dans une longue série tous les événements tendront à arriver avec une fréquence relative proportionnelle à leurs probabilités objectives »[71]. Venn s'élève aussitôt contre cette « tendance invétérée à objectifier nos conceptions, même là où ces conceptions n'ont aucun droit à exister » (ibidem). L'auteur s'élève ici contre l'interprétation objectiviste classique.

Un autre auteur énonce ainsi le lien entre probabilités a priori et fréquences :

« Nous nous étions proposés de déduire de la notion subjective de probabilité des conséquences vérifiables expérimentalement. Le théorème de Bernoulli répond à cet objet. Grâce *à* ce théorème, la probabilité apparaît comme absolument assimilable à une grandeur physique, dont une mesure expérimentale donne une valeur approchée. La série de n expériences effectuées dans

les mêmes conditions nous permet de mesurer la probabilité, et la fréquence est le résultat de la mesure. »[72] Or Richard von Mises s'élève lui aussi contre l'existence d'un tel lien entre la probabilité subjective et l'expérience. Il est nécessaire de rentrer quelque peu ici dans les détails.

Notons que l'expression de loi des grands nombres vient de Poisson. Il appelle ainsi deux choses distinctes : la première est énoncée dans l'introduction des *Recherches sur la probabilité des jugements*, et pourrait être formulée ainsi : « Dans différents domaines, les phénomènes empiriques semblent obéir à une certaine loi générale, qu'on pourrait appeler la Loi des Grands Nombres. Cette loi stipule que les rapports de nombres dérivés de l'observation d'un très grand nombre d'événements similaires restent pratiquement constants, pourvu que ces événements soient gouvernés en partie par des facteurs constants, et en partie par des facteurs variables dont les variations sont irrégu1ières et ne causent aucun changement systématique dans une direction définie.

La deuxième est une généralisation de la loi faible des grands nombres lorsque les probabilités ne sont pas constantes[74]. Or ces deux énoncés ne sont pas équivalents selon von Mises. Le premier est une autre manière de formuler le premier axiome de sa théorie, celui de l'existence d'une limite des fréquences relatives dans une séquence infinie.

Le second ne concerne que certains types de séquences, qu'il appelle séquences de Bernoulli. Il montre ainsi qu'il existe des séquences pour lesquelles le premier axiome de la théorie "marche", alors que le théorème de Poisson ne marche pas (ibidem).

Si Poisson considère que la loi empirique qu'il formule dans son introduction est équivalente à son théorème mathématique, c'est parce qu'il comprend sous la notion de probabilités deux choses différentes. « Au début de ses calculs, ...il voulait dire le rapport des cas favorables à l'ensemble de tous les cas possibles. Cependant, il interprète à la fin de ses calculs l'expression "presque 1" dans un sens qui veut dire que l'événement arrivera dans presque tous les jeux[75] ». Or ce changement de sens ne semble pas admissible à von Mises (ibidem). C'est pourquoi il précise les relations qu'ont ces divers théorèmes mathématiques avec les axiomes de sa théorie.

Ses deux axiomes (l'existence de la limite de fréquence relative, et l'axiome du hasard) sont en fait nécessaires pour la déduction des théorèmes mathématiques, car les différents théorèmes disent plus sur le comportement de séries d'épreuves que la simple admission de l'existence de limites des fréquences relatives (p.110). Chez von Mises les séquences de Bernoulli sont des conséquences de l'axiome du hasard, mais elles ne constituent pas en elles-mêmes un substitut à l'exigence de fortuité. Le théorème de Bernoulli-Poisson-Tchebyschev n'exprime qu'un type de hasard (p. 113).

Si tous les collectifs sont des séquences de Bernoulli, en revanche toutes les séquences de Bernoulli ne sont pas des collectifs (ibidem).

La confusion repose sur l'identification de deux notions de fréquences. La première qui est utilisée dans la définition de la probabilité veut dire simplement le rapport de nombres d'occurrences d'un attribut (par exemple le nombre six) dans une série infinie d'observations ou d'épreuves. La seconde est plus précise : elle veut dire la fréquence relative de l'événement ($\left|\frac{Sn}{n-p}\right| < e$) dans une série de n épreuves.

Dans le cas du théorème de Bernoulli chaque groupe de n épreuves est considéré comme un élément d'un collectif. Et c'est dans ce collectif qu'on peut démontrer le théorème de Bernoulli (p.86-87).

Ainsi, si on prend les tables des racines carrées d'un nombre, et qu'on considère la sixième décimale, et qu'on la remplace par un zéro si c'est un chiffre pair et par un 1 si c'est un chiffre impair, la fréquence des zéros et des uns tend vers **1/2.** En revanche le théorème de Bernoulli n'est pas valable pour de telles séquences car on trouve des groupes dont le nombres de 0 ou de 1 est de plus en plus croissant[76].

Cet exemple, et d'autres similaires prouvent qu'il existe des séries pour lesquelles il existe une limite des fréquences relatives, mais pour lesquelles le théorème de Bernoulli n'est pas valide. Par conséquent le théorème de Bernoulli ne peut fournir un pont entre la définition classique de la probabilité, et la notion de fréquence relative. Bien évidemment, la série considérée ci-dessus n'obéit pas à l'exigence de l'axiome de fortuité. Nous allons examiner brièvement dans la section suivante en quoi consiste cet axiome.

§4. L'apport positif et théorique de von Mises au calcul des probabilités repose sur la notion de collectif. Le collectif est une séquence infinie de résultats qui doit satisfaire à deux préréquisits : 1) l'existence d'une limite des fréquences relatives, et 2) l'axiome du hasard ou de fortuité, ou encore d'irrégularité.

Ce deuxième axiome a été beaucoup discuté dans la littérature. Intuitivement von Mises le présente ainsi. Après une sélection des éléments d'une suite, la limite de la fréquence

relative d'un attribut doit rester la même dans la suite obtenue. Ou présenté autrement, il n'existe aucun système de jeu qui soit toujours gagnant. On ne peut pas changer les chances d'occurrences d'un événement par une sélection quelconque des éléments de la série (p.24). « Les valeurs limites des fréquences relatives doivent être indépendantes de tous les choix de sélection possibles » (p.25). Un attribut doit avoir la même fréquence relative dans le nouveau collectif obtenu par un "choix de sélection" du premier.

§5. On a objecté à cette théorie qu'elle n'était pas exempte de contradiction, étant donné qu'on peut choisir des éléments de la suite de telle manière que la fréquence obtenue d'un attribut soit zéro ou un.

Je vais examiner dans cette section quelques-unes de ces critiques, en m'inspirant principalement de celles qui sont formulées par Maurice Fréchet, et par de Finetti. Le premier représente une tendance objectiviste, dans laquelle la probabilité serait comparable à une grandeur physique que les fréquences relatives mesurent de manière plus ou moins précise[77]. Cette tendance était partagée par un certain nombre de mathématiciens français, parmi lesquels se trouvent Paul Lévy, ou Emile Borel.

Ces mathématiciens, en particulier Emile Borel sont partis de la définition classique, et l'ont améliorée. Ils ont introduit de nouveaux concepts mathématiques (théorie de la mesure de Lebesgue) pour traiter de quelques paradoxes et étendre la théorie aux ensembles dénombrab1 es. Dans cette théorie, la probabilité devient une fonction d'ensemble additive et non négative. Ce nouveau cadre conceptuel, Jerzy Neyman l'a appelé théorie classique modernisée. C'est de ce cadre que les critiques de Maurice Fréchet prennent sens.

La première critique que Maurice Fréchet porte contre la théorie de von Mises, est que celle-ci exclut du calcul des probabilités un certain nombre de suites qui ont une probabilité dans la théorie de la classique modernisée, alors qu'elle en admet d'autres qui n'ont pas de sens (par exemple des suites qui convergent par valeurs supérieures, ce qui contredit le caractère oscillant des suites fortuites dans la théorie classique

modernisée). « Pour nous, l'hypothèse qu'on examine seulement les suites d'épreuves où la fréquence converge, est celle qui choquait le plus. »[78] D'autre part elle ne rend pas compte des suites finies. Enfin elle veut donner un modèle arithmétique du hasard, ce qui est contradictoire, car l'esprit humain ne peut imiter le hasard, selon une remarque d'Emile Borel.

§6. Il est difficile pour un profane de prendre position dans ce débat. On sait que von Mises a accepté les amendements de sa théorie apportée par Wald, et qui consiste à limiter l'ensemble de tous les choix de position possibles, en le donnant la puissance du dénombrable. Il montre ainsi qu'il existe un collectif répondant aux réquisits. Pour plus de détails, on peut se rapporter à l'examen critique détaillé que fait Fine des conceptions fréquentistes, au chapitre IV de son livre. Ce qu'on retient de l'examen, c'est que la théorie du collectif de von Mises améliorée par Wald, Ville, devient non contradictoire, mais au prix d'une relativisation de la notion de collectif.

Si on se place du point de vue statistique, enfin, et qu'on se demande quel est le rôle et l'utilité des fréquences dans le domaine de l'estimation, peut-on dire que les fréquentistes ne recourent qu'à l'expérience ? Cette question est posée par Fine. Si on part de l'exigence minimale que doit satisfaire un estimateur (qui une fonction des fréquences observées), et si on suppose la probabilité fixe (Fine p.100.) alors la probabilité que l'écart entre p et son estimateur soit inférieure à une grandeur donnée est soit zéro, soit un. Et si p est une variable aléatoire, avec une distribution a priori, il existe des estimateurs fonctions des fréquences qui satisfont à l'exigence minimale (ibidem). Mais comme la distribution a priori, ou bien ne dépend pas des données, ou alors dépend d'autres données que les fréquences, l'estimateur en question ne dépend plus exclusivement des fréquences. Dans les deux cas les fréquences relatives seules ne mènent pas à des mesures de probabilités.

Ces difficultés, et d'autres encore exposées par Fine, n'obligent-elles pas à prendre un point de vue plus cohérent sur le concept de probabilité en adoptant un point de vue subjectif ? C'est ce à quoi invite l'approche personaliste de de Finetti, que je vais examiner brièvement.

CHAPITRE V PROBABILITES PERSONNELLES

§1. Ce dernier chapitre sera consacré à la conception développée par de Finetti et que Léonard Savage a appelé personnalisme. Je me baserai essentiellement sur l'important mémoire de 1937. « La prévision : ses lois logiques, ses sources subjectives », ainsi que sur le livre, *Theory of Probability, a critical introductory treatment*.

Quelles sont les conceptions philosophiques qui inspirent l'auteur et qui déterminent son approche du concept de probabilité ?

L'auteur avoue avoir été influencé par Locke, Hume, Berkeley, et plus près de nous par le pragmatisme et l'opérationalisme (Mach, Bridgman).

Pour lui la différence qui oppose subjectivistes et objectivistes n'est pas d'ordre philosophique, mais méthodologique (p.6). Ce souci de ramener les querelles principielles à des querelles de méthodes le conduit à redéfinir les concepts utilisés jusque là dans la littérature.

Par exemple la notion d'indépendance. D'un point de vue objectif deux événements sont indépendants, si l'arrivée de l'un n'influence pas l'arrivée de l'autre. Du point de vue subjectif deux événements seront indépendants si la connaissance du résultat de l'un ne change pas l'évaluation de probabilité de l'autre (p.7).

De même il est important pour l'auteur de démontrer que tous les théorèmes ou axiomes du calcul des probabilités peuvent découler d'une conception subjective.

Les considérations intuitives qui servent de base au calcul sont les suivantes : on peut établir une échelle de comparaison entre ce qui parait incertain, certain et impossible, à condition que l'on admette qu'un événement est soit aussi probable, soit plus probable, soit moins probable qu'un autre. On définit ainsi une relation d'ordre sur l'ensemble des événements incertains. Cet ordre est compatible avec l'addition logique. Ces relations permettent de déduire l'axiome des probabilités totales[80].

On peut donner aussi une évaluation qualitative, directe, du degré de probabilité qu'un individu accorde à un événement déterminé, en se fondant sur la notion de pari. Un pari est le prix p qu'un individu est prêt à payer pour échanger une somme S (positive ou négative) qui dépend de l'arrivée d'un événement E contre une somme pS. Ce nombre p est la mesure du degré de probabilité attribué par l'individu en question à l'événement E (ibidem).

D'autre part, pour lever toute ambiguïté de Finetti précise qu'il entend événement comme un fait singulier. Lorsqu'il y a plusieurs épreuves, il parle d'épreuves d'un même phénomène, et non pas d'un même événement. Une épreuve constitue en elle-même un événement.

Le pari que fait un individu doit être cohérent, exempt de contradiction. Je ne peux pas parier avec lui en étant assuré de gagner à coup sûr. Cette notion de cohérence est extrêmement importante, car elle constitue le principe sur De même il est

important pour l'auteur de démontrer que tous les théorèmes ou axiomes du calcul des probabilités peuvent découler d'une conception subjective.

Les considérations intuitives qui servent de base au calcul sont les suivantes : on peut établir une échelle de comparaison entre ce qui parait incertain, certain et impossible, à condition que l'on admette qu'un événement est soit aussi probable, soit plus probable, soit moins probable qu'un autre. On définit ainsi une relation d'ordre sur l'ensemble des événements incertains. Cet ordre est compatible avec l'addition logique. Ces relations permettent de déduire l'axiome des probabilités totales'".

On peut donner aussi une évaluation qualitative, directe, du degré de probabilité qu'un individu accorde à un événement déterminé, en se fondant sur la notion de pari. Un pari est le prix p qu'un individu est prêt a payer pour échanger une somme S (positive ou négative) qui dépend de l'arrivée d'un événement E contre une somme pS. Ce nombre p est la mesure du degré de probabilité attribué par l'individu en question a l'événement E (ibidem). D'autre part, pour lever toute ambiguïté. de Finetti précise qu'il entend événement comme un fait singulier. Lorsqu'il y a plusieurs épreuves, il parle d'épreuves d'un même phénomène, et non pas d'un même événement. Une épreuve constitue en elle-même un événement.

Le pari que fait un individu doit être cohérent, exempt de contradiction. Je ne peux pas parier avec lui en étant assuré de gagner à coup sûr. Cette notion de cohérence est extrêmement importante, car elle constitue le principe sur lequel est fondé tout le calcul. « Ce calcul apparaît alors comme

l'ensemble des règles auxquelles l'évaluation subjective des probabilités de divers événements par un même individu doit être assujettie si l'on ne veut pas qu'il y ait entre elles une contradiction fondamentale »[81].

L'objet du calcul des probabilités n'est donc plus une seule fonction P(E) des événements E, mais l'ensemble de toutes les fonctions P(E) correspondant à des opinions admissibles (p.9).

§2. Une fois admis le point de vue subjectif, il faut préciser les relations entre probabilités et fréquences.

Cette analyse se divise en deux (p.18). Une partie consacrée aux rapports entre évaluations de probabilités et prévisions de fréquences futures. Une autre consacrée aux relations entre l'observation de fréquences passées et la prévision de fréquences futures.

La première partie est simple et la seconde plus délicate. Elle repose sur un théorème : soit une famille d'événements quelconque {E„} de probabilités $p_1, p_2, \ldots, p_n$., et soient a_n les probabilités que zéro, ou un seul, ou deux, ou n de ces événements se réalisent. Alors la moyenne arithmétique des p_i est égale à l'espérance mathématique de la fréquence, définie comme la variable aléatoire qui prend les valeurs 0/n, 1/n, 2/n, n/n selon que 0,1, 2, … n des E_n se réalisent.

Un exemple montrera plus facilement le contenu de ce théorème. Supposons qu'on cherche à évaluer la

probabilité qu'un individu donné A meure dans l'année (p.20). On considère pour cela d'autres événements, par exemple le décès pendant une année d'individus du même âge que Monsieur A et vivant dans le même pays. Supposons que sur 1000 individus, on prévoit que **13** vont mourir. La probabilité du décès de Monsieur A sera alors de **0,013**. Si de plus, je suis un médecin qui sait que Monsieur A est atteint d'une maladie incurable, alors cette probabilité sera plus près de **1** que de cette valeur moyenne.

Le processus d'évaluation comprend trois phases : le choix d'une classe d'événements comprenant celui que l'on considère ; la prévision de la fréquence ; la comparaison entre la probabilité des événements singuliers et l'évaluation de la fréquence. On peut remarquer que déjà le choix d'une classe d'événements est en soi arbitraire. De même toute classification rigide des événements en classes ou en séries chez les fréquentistes ne peut prendre qu'une valeur arbitraire.

En conclusion : « Ainsi le critère basé sur la notion de fréquence se réduit, comme celui des événements également probables, à une méthode pratique pour ramener certaines évaluations subjectives de probabilités è d'autres évaluations, elles-mêmes subjectives, mais préférables, soit parce que plus accessibles à une estimation directe, soit parce qu'une estimation plus grossière ou même de nature purement qualitative suffit pour les conclusions envisagées. »[82]

Maintenant passons à la question plus délicate de l'influence des événements passés sur la prévision des événements futurs. On sait que Laplace avait traité cette question en utilisant le cadre formel du théorème de Bayes.

De Finetti ne fait pas autre chose. En regardant la définition de la p.26 le lecteur retrouve la définition de la probabilité conditionnelle. C'est l'interprétation qui change. En effet, il n'y a pas une différence de nature entre la probabilité a posteriori et la probabilité a priori. La première n'est pas indépendante de la seconde et ne la remplace pas. Elle découle du même jugement a priori « en en retranchant seulement, pour ainsi dire, les composantes du doute relatives aux épreuves dont le résultat a été acquis ». Les termes a priori et a posteriori ne veulent plus dire que simplement avant et après l'information.

La règle de la probabilité conditionnelle et le théorème de Bayes qui en découle permettent de rendre compte du processus d'apprentissage par l'expérience, et explique le phénomène de "dépendance stochastique à travers une augmentation de l'information". L'expérience modifie une opinion préexistante à la lumière d'une situation nouvelle. Mais il faut être attentif : si nous modifions une opinion préexistante, nous ne la corrigeons cependant pas parce que cela impliquerait l'existence d'une probabilité réelle, qui bien évidemment n'existe pas.

Ce processus n'est valable que si toutefois on ait pris soin de ne pas isoler notre opinion par une supposition d'indépendance. Car sous cette hypothèse, notre opinion ne peut plus être modifiée par des observations ultérieures (qui constituent autant de sources d'information).

« Dans le cas de l'indépendance le problème de l'induction n'existe pas […] Toute possibilité d'apprentissage est exclue - écartée par le principe de contradiction »[84].

Pour rendre compte du processus d'induction dans le problème qui nous occupe, c'est-à-dire le passage de fréquences observées à la prévision des événements futurs (puisqu'en fait c'est un problème d'induction), il faut une condition plus faible que l'indépendance. Cette condition, c'est l'échangeabilité.

La notion a été introduite dès les années 1930 et est amplement développée dans l'article de 1937. Cette notion s'applique aussi bien à des événements qu'à des variables aléatoires. En 1937 de Finetti appelle « équivalence » ce qu'il appellera plus tard, sur une suggestion de Polya semble-t-il, « échangeabilité ». Cette notion veut dire que l'ordre d'occurrence des événements est non-pertinent pour la prévision des fréquences futures.

Montrons cela sur un exemple concret. Etant donné n événements, la probabilité de l'arrivée de h d'entre eux sera notée $\alpha_n(n)$, avec h **= 0,1,2,** … n. Bien évidemment $\sum\alpha_n(n)=1$.

Ces probabilités seront quelconques (à condition de remplir la notion de cohérence bien sûr). On n'a plus la distribution binomiale qu'on a habituellement dans la supposition d'indépendance, mais une distribution dans laquelle tous les produits des n événements ont la même probabilité. En effet les probabilités que nous avons définies ci-dessus ne sont que les sommes des combinaisons de h objets pris parmi n. Leur

moyenne est égale au α_n divisée par les combinaisons de r objets parmi n. Or si on considère l'ordre comme indifférent, tous les résultats comportant une même fréquence r/n, ont la même probabilité, à savoir celle définie ci-dessus (la moyenne des α_n). On retrouve le résultat considéré dans la première partie (la moyenne des probabilités est égale à l'espérance des fréquences. Cf.p.63). Donc lorsque l'ordre des épreuves n'intervient pas dans nos jugements de probabilités, les événements seront dits équivalents ou échangeables par rapport à nos jugements de probabilité. Dans ce cas la probabilité d'un événement quelconque E subordonné à l'hypothèse A d'un résultat respectivement favorable et défavorable dans r et s épreuves déterminées, ne dépend pas de la nature des événements choisis, mais simplement des nombres r et s comme on peut le voir en appliquant la définition de la probabilité conditionnelle (démonstration p.29).
De Finetti démontre aussi que les lois des grands nombres sont aussi valables pour les événements échangeables.

Enfin, tous les cas où les conditions d'échangeabi1ité sont définies pourront être appliqués aux épreuves dans lesquelles on cherche une « probabilité constante, mais inconnue » (schéma des urnes, urnes de Polya).

§2. Cette notion d'échangeabilité permet de résoudre de manière naturelle (c'est à dire sans hypothèses supplémentaires, ou adhockeries) des problèmes d'inférences statistique.

Le problème de l'inférence statistique est posé différemment par Fisher et par Neyman et Pearson.

Chez Ficher il s'agit de construire un test d'hypothèses dans lequel nous évaluons la crédibilité d'une hypothèse H_0 à la lumière des faits. Ces faits permettent de rejeter l'hypothèse nulle, mais pas de la confirmer. Fisher pense que les tests d'hypothèses caractérisent la démarche scientifique.

L'approche de Neyman-Pearson de l'inférence statistique est plus pragmatique. Un test d'hypothèses conduit à une décision, qui est susceptible de deux types d'erreur (ou bien on accepte une hypothèse fausse, ou bien on rejette une hypothèse vraie). Il faut minimiser le risque. Mais les deux stratégies sont contraires. Cette différence d'approches des tests d'hypothèses a donné lieu à des controverses acerbes que nous ne résumons pas.

Par rapport à ces deux approches objectivistes, l'inférence bayésienne offre une approche unifiée de l'inférence statistique. De Finetti donne quelques exemples de distributions a posteriori dans le chapitre 12 de *Theory of probability*, selon que l'on connaisse ou non la moyenne ou la variante, et selon les hypothèses que l'on fait sur la distribution initiale. Lindley montre aussi qu'on peut construire des intervalles de confiance bayésiens.

CONCLUSION

Il y aurait d'autres questions annexes connectées à l'approche bayésienne, notamment la notion d'utilité et la prise de décision. Cependant les traiter, même schématiquement, nous éloignerait un peu trop de notre objet principal.

De l'autre côté, on pourrait aussi chercher dans la physique, si on a une position objectiviste, d'autres fondements que la fréquence, à la notion de probabilité, comme par exemple la notion de propension. Mais ces raffinements exigeraient un plus long développement.

Comme ultime question, on peut se demander si l'opposition fréquentistes/bayésiens recouvre l'opposition objectivistes/subjectivistes. Il n'en est pas ainsi, parce que le premier couple d'opposition a une extension plus petite que le second. Mais une raison plus profonde existerait, comme le suggérait L. Savage. Les fréquentistes chasseraient les procédures subjectives par la porte pour les réintroduire par la fenêtre. Far rapport aux différentes propriétés que recherche un fréquentiste (absence de biais, erreur quadratique minimale, niveau de confiance…), l'approche bayésienne offre une vue plus unifiée de l'inférence statistique. Quant au concept de probabilité lui-même, en peut reconnaître qu'il a une pluralité de sens qu'il serait difficile de réduire à l'analyse. Mais si une réduction peut être tentée, nous dirions que les sens subjectifs prendraient le dessus sur les sens objectifs.

NOTES

1 Laplace, *Essai philosophique sur les probabilités*, Christian Bourgeois éditeur, avec une préface de René Thom et une postface de B. Bru, Paris, 1966, p.32-33.
2 *Les principes physiques de la théorie des quanta*, trad. française de : *Die Physikalischen Prinzipien der Quantentheorie* (1930), Paris, 1972, nouveau tirage, p.15.
3 Cf. Feynman, Leighton, Sands, *Lectures on Physics*, T.III chap.1
4 Cf. pour plus de détail la discussion des expériences in Heisenberg, op.cit. p.15 sqq.
5 Physical Review, t.47,1935, p.777
6 Lettre d'Einstein à Karl Popper, in *La logique de la découverte scientifique*, trad.fr. p. 471.
7 *La nature de la physique*, p.172
8 *Theory of probability*, préface, p. x
9 *Logique de la découverte scientifique*, trad. fr. , §48, p. 150
10 Brochard, p.108
11 *Académiques*, XXIII, 72 sqq
12 Sextus Empiricus, *Contre les logiciens*, I, 159
13 Diogène Laërce, VII, 46
14 Sextus Empiricus, op.cit., I, 372
15 Sextus, idem, I, 227
16 Sextus, idem, I, 252
17 Cf. aussi *Académiques*, II, vi, 18 et xxiv, 77: « Ce qui est compris est une représentation (c'est le mot que nous avons usé hier pour rendre *phantasia*)imprimée et formée d'après l'objet d'où elle provenait, telle qu'elle ne pourrait être si cet objet n'existait pas. » (trad. Emile Bréhier, in Pléiade, p.196).
18 Emile Bréhier, *Chrysippe et l'ancien Stoïcisme*, p.97.
19 Cicéron, *Académiques*, II, xxvi, 83. Nous suivons en outre l'exposé de Victor Brochard, *Les sceptiques grecs*, Paris, 1887, réimprimé par Vrin, nouv. éd., 1986.
20 *Académiques*, II, xxviii, 91.
21 Sextus Empiricus, *Contre les Logiciens*, I, 156-184.
22 *Académiques*, II, xviii, 60.
23 *Académiques*, xii, 36.
24 Ore p.193.
25 Ore [1953] p.145. •
26 Ore, *op.cit.*; Maistrov, p. 19 sq.
27 David, *op. cit.*,p.56.
28 Hacking 1975 p.53.
29 Hacking, op cit., p.52 et David, p. 192.
30 Cf. Feller [1968].
31 Pour ce passage comparer Maistrov op. cit. p.30 et Hacking p.53.
32 Cf. Ore 1960 p.414.
33 Cf. Kendall, in Pearson & Kendall 1970.
34 Poirier p.11.
35 *L'art de conjecturer*, traduit du latin de Jacques Bernoulli par L.G.F.Vastel, p.11.
36 *Essai philosophique*, postface, p.250.
37 Recherches, 1°, sur l'intégration des équations différentielles aux différences finies, et sur leur usage dans la théorie des hasards. 2°, sur le principe de la gravitation universelle, et sur les inégalités séculaires qui en dépendent. 1776, O.C., VIII, 69-275.
38 pages 144 à 153 dans l'édition citée.
39 op. cit., p. 145.
40 idem, p.144.

41 Cf. *Théorie analytique*...dans l'édition de l'Imprimerie Royale, Paris, 1847, p.197. 42 *Mémoire sur les probabilités des causes par les événements*, p.29.
43 Stigler, p.102, et *Théorie analytique*...p.198.
44 *Théorie analytique*...p.199.
45 *Mémoire sur la probabilité des causes par les événements*, p.33.
46 Voir Stigler, p.132.
47 Sheynin, p.155.
48 *Mémoire sur les probabilités*, p.419.
49 Comparer les pages 30 à 33 du mémoire de 1774 et les pages 419 à 422 du mémoire de 1781.
50 *Mémoire sur les probabilités*, ibidem, p.423.
51 Probability, Statistics and Truth, p.I17 sqq.
52 Mémoire sur les probabilités, p.431-432.
53 *Théorie analytique*..., édition de l'Imprimerie Royale, p. 308.
54 Pour une discussion de ce point cf. l'article de Sheynin: Laplace's Work on probability.
55 *Essai philosophique*, édition citée, p.34. 56
Pour ce paradoxe, cf. Poincaré, *La Science et l'Hypothèse*. éd. Flammarion, p.I92.
56 Mémoire sur les probabilités, t. IX, O.C., p.384-385.
57 Loc. cit.
58 Cf. *De Interpretatione*, 13, 23a10.
59 *Mémoire sur les probabilités*, p.385.
60 Cf. *Le Hasard*. "Opere cit., p.54
61 *Mémoire sur la probabilité des causes par les événements*, p. 56.
62 ibidem, p.6
63 *Logic of chance*, pp.4-5.
64 Hodges & Lehmann, p.10.
65 Probabilités, Hasards, Inductions, p.83.
66 *Probability, Statistics and truth*, p.67.
67 idem, p.70.
68 idem, p.78.
69 *Logic of Chance*, p.91.
70 Paul lévy, RMM, 1963, p.29-30.
71 Cité d'après von Mises, op. cit., p.104-5. Voir aussi Stigler p.185.
72 ibidem. "Von Mises, op.cit., p.109.
73 Von Mise, op.cit., p.111 sqq.
74 Jean Ville, p.9: « La probabilité d'un événement, dans une catégorie d'épreuves déterminée, est une constante physique, dépendant de l'événement et de la catégorie d'épreuves, dont on obtient une valeur empirique en procédant à un grand nombre d'épreuves indépendantes et en notant la fréquence de l'événement ».
75 Fréchet, p.26.
76 *Theory of probability*, Préface, p.xii. et p.70.
77 de Finetti 1937, p.5.
78 op.cit., p.7.
79 de Finetti 1937, p.22.
80 Cf. aussi *Theory of Probability*, t.2, p.199 : "The vital element in the inductive process, and the key ta every constructive activity of the human mind, is then seen to be Bayes's theorem"
81 *Theory of Probability*, p.206.

REFERENCES BIBLIOGRAPHIQUES.

BAYES Thomas [1763], "An essay towards solving a problem of chance", in Philosoph1ca1 Transactions of the Royal Society of London. **53**. 370-416, London. Reprint in Pearson & Kendall [1970], 131-156.

BERTRAND Joseph [1889]. *Calcul des probabilités,* Paris. Reprint Chelsea Publishing Company, New York, 1972.

BOREL Emile [1909]. *Eléments de la théorie des probabilités*. 3e édition, Paris. Hermann.

BOREL Emile [1913]. *Le hasard*, 2e édition. Paris. Alcan, 1932.

BREHIER Emile [1910]. *Chrysippe et l'ancien Stoïcisme*. Paris, PUF. Deuxième édition revue. 1951, PUF.

BROCHARD Victor [1887], *Les Sceptiques grecs,* Paris, Reprise Vrin, 1965.

CAVA1LLES Jean [1940] "Du collectif au pari. A propos de quelques théories récentes sur les probabilités". Revue de Métaphysique et de Morale. **47**, 139-163.

CICERON [1933], *Academica*, with an english translation by H. Rackham. London, Cambridge. W.Heinemann Ltd and Harvard University Press, 1961.

COURNOT Antoine-Augustin [1843]. *Exposition de la théorie des chances et des probabilités*. Paris, Hachette. Réédité par Bernard BRU, in Œuvres complètes, t.1, Paris. Vrin-Cnrs. 1984.

DAVID Florence N. [1952]. *Games, Gods and Gambling. The origins and history of probability and statistical ideas from the earliest times to the Newtonian era.* London, Griffin & Co.

DE MOIVRE Abraham [1718], *The doctrine of chances, or a method of calculating the probabilities of events in Play,* London, 2nd edition, 1738; H. Woodfall 3rd edition 1756, London, Miller.

DEMAN Thomas [1936]. Article « Probabilisme » in *Dictionnaire de Théologie Catholique*, t.XIII, col.417-519, Paris, Letousé et Ané.

DIOGENE LAERCE [IVe siècle après J.C.], *Vies, doctrines et sentences des philosophes illustres*. Paris, Garnier Frères. 1965.

DUMONT Jean-Paul [1965], *Le scepticisme et le phénomène*, Seconde édition, Paris, Vrin.

FELLER William **[1968].** *An introduction to Probability Theory and Its Applications*, t. 1, 3rd edition, New York. Wiley & Sons.

FEYNMAN Richard [1965], *La nature de la physique*, traduit de l'américain par H. Isaac, J.M. Lévy-Leblond et F. Balibar, Paris, Seuil.

FINE Terrence L. [1973] *Theories of probability, an examination of foundations*, New York. Academic Press.

FINETTI Bruno de [1937], "La prévision : ses lois logiques, ses sources subjectives", in Annales de l'Institut Henri Poincaré. **7**, 1-68, Paris.

FINETTI Bruno de [1939]. "Compte rendu critique du colloque de Genève sur la théorie des probabilités", in Actualités scientifiques et industrielles n°766. Paris, Hermann.

FINETTI Bruno de [1974]. *Theory of probability, A critical introductory treatment*, translated by A. Machi &A. Smith, 2 volumes. New York, Wiley.

FISHER Ronald A. [1922] "The mathematical foundations of theoretical statistics". Philosophical Transactions of the Royal Society. A, 222, 309-368.

FISHER Ronald A. [1935], "The logic of statistical inference". Journal of the royal statistical society, 98, 39-54.

FISHER Ronald A. [1955], "Statistical method and scientific induction", Journal of the royal statistical society, 5, **17**. 69-78.

FRECHET Maurice [1939], "Exposé et discussion de quelques recherches récentes sur le fondement du calcul des probabilités", in Actes du colloque de Genève sur la théorie des probabilités, pp.23-55, Paris, Hermann.

GIGERENZER Gerd et alii [1989], The Empire of Chance. How probability changed everyday life. Cambridge, Cambridge university Press.

HACKING Ian [1971], "James Bernoulli's Art of Conjecturing", British Journal of philosophy of science, 22, 209-229. London.

HACKING Ian [1975], *The Emergence of Probability, A phiiosophical study of early ideas about probability, induction and statistical inference*. Cambridge university Press.

HEISENBERG Werner [1932]. *Les principes physiques de la théorie des quanta*, Paris. Gauthier-Villars. 1972 (nouveau tirage).

HODGES J.L. Jr and LEHMANN E.L. [1970], *Basic concepts of probability and statistics*, 2nd ed., San Francisco, Holden-day.

KEYNES J.M. [1921]. *A treatise on probability*. London, Macmillan, 1973, 3rd ed.

KOLMOGOROV A.N. [1933]. *Foundations of the Theory of Probability*, New York, Chelsea Publishing Company. 1956.

LAPLACE Pierre-Simon [1774], "Mémoire sur la probabilité des causes par les événements", in *Œuvres Complètes*, t. VIII, 27-65, Paris, Gauthier-Villars, 1891.

LAPLACE Pierre-Simon [1781], "Mémoire sur les probabilités". in *Œuvres Complètes*, t. IX, 363-465, Paris, Gauthier-Villars. 1891.

LAPLACE Pierre-Simon [1814], *Essai philosophique sur les probabilités*, Paris, Christian Bourgeois, 1986.

LAPLACE Pierre-Simon [1820], "Théorie analytique des probabilités", 3ième édition avec suppléments, in *Œuvres Complètes*, t. VII. Paris. Gauthier-Villars. 1886.

LARGEAULT Jean [1979], *Hasards, Probabilités, Inductions*, Toulouse, Presse de l'Université de Toulouse-le-Mirail 1. série A, tome 42.

LEIBNIZ Gottfried Wilhelm, *Nouveaux Essais sur l'entendement humain*, édités par J. Brunschwig. Nouveau tirage, Paris, Flammarion, 1990.

LEVY Paul [1963], "Les fondements du calcul des probabilités" Revue de Métaphysique et de Morale, **68**, 25-56.

MAISTROV L.E. [1974], *Probability theory, an historical Sketch*, translated from the Russian by S. Kotz, New-York, Wiley.

MISES Richard von [1957], *Probability, Statistics and Truth*. second revised English edition prepared by Hilda Geiringer. London, Allan & Unwin, New York, Macmi11an.

MOLINA Edward C. [1930]. "The theory of probability. some comments on Laplace's Théorie analytique". Bulletin of the American mathematica1 society. **36**, 369-392.

MOLINA Edward C. [1931]. "Bayes's theorem. An expository Presentation", Annals of mathematical statistics. 2. 23-37.

NEYMAN Jerzy [1937], "Outline of a theory of statistical estimation based on the classical theory of probability, Philosophical Transactions. A, 236, 333-380, London.

ORE Oystein [1953], Cardano, the Gambling scholar, with a translation from the Latin of Cardano's Book on games of chance, by S.H. Gould, New York, Dover Publications inc.

ORE Oystein [1960], "Pascal and the invention of probability theory". American Mathematica1 Monthly. **67**, 409-419.

PASCAL Blaise [1962], Œuvres Complètes, éditées par LAFUMA, Paris, Seuil.

PEARSON E.S. &. KENDALL M.G. [1970], Studies in history of statistics and probability. London, Griffin.

PEARSON Karl [1924],"Historical note on the origin of the normal curve of errors", Biometrika, **16**, 402-404, London.

PEARSON Karl [1925], "James Bernoulli's Theorem". Biometrika, **17**, 201-210. London.

PEARSON Karl [1976]. The history of probability in the 17th and 18th centuries, against the changing background of inte1iectual scientific and religious thought. Lectures by Karl Pearson giving at University College during the academic sessions 1921-1933, edited by E.S. Pearson, London, Griffin.

POINCARE Henri [1909], *La science et l'hypothèse*. Paris. F1ammarion.

POINCARE Henri [1912]. Calcul des probabi1ités. 2nd édition. Paris, Gauthier-Villard. Réimpression J. Gabay, 1987.

POIRIER René [1931]. Remarques sur la probabilité ces inductions, Paris, Vrin.

POISSON Siméon-Denis [1837], Recherches sur la probabilité des jugements en matière criminelle et en matière civile, précédées des règles générales du calcul des probabi1ités. Paris, Bachelier.

POPPER Karl [1973], La logique de la découverte scientifique, traduit de l'anglais par N. Thysssen-Rutten et P. Devaux, Préface de J. Monod, Paris, Payot.

ROBIN Léon [1944] *Pyrrhon et le scepticisme grec,* Paris, P.U.F.

SAVAGE Léonard J. [1961], "The foundations of statistics reconsidered", in Proceedings of the fourth Berkeley symposium on mathematical statistics and probability, vol.1 pp. 575-586, Berkeley, University of California Press.

SEXTUS EMP1RICUS [1935], *Against the logicians*, Loeb Classical Library, Cambridge, Mass., Harvard University Press.

SHEYNIN Oscar B. [1968], "On the early history of the law of large numbers", Biometrika, 55, 453-467.

SHEYNIN Oscar B. [1978], "Poisson's Work in Probability", Archive for history of exact sciences, **18**, 245-300, Springer Verlag.

STIGLER M. Stephen [1982], "Thomas Bayes's Bayesian inference" Journal of the Royal Statistical Society, A, 145, 250-258.

STIGLER M. Stephen [1986], The history of statistics, The Measurement of Uncertainty before 1900. Cambridge, Mass., The Belknap Press of Harvard University Press.

TODHUNTER Isaac [1865], A history of the mathematical theory of probability, from the time of Pascal to that of Laplace, London and Cambridge. Reprint Chelsea Publishing Company, New York, 1965.

VENN John [1888], Logic of chance, 4th ed., New York, Chelsea, Reprint of the 2nd edition, 1962.

VILLE Jean [1939], Etude critique de la notion de collectif. Paris, Gauthier-Villars.